

```
QY 362 TTGACGAATTAATGACAACTAGTGGCTTTAAAGTCATCAGCATGATGACAGAGAG 421
XX |||
XX |||
DB 506 TTAGCAGAAATTAATGACAACTAGTGGCTTTAAAGTCATCAGCATGATGACAGAGAG 565
QY 422 GAGTCCCATTTACAGTATCCGAGAGCTTCTCTCTGAAAGGGTTTGAACATGCCAAAT 481
DB 566 GAGTCCCATTTACAGTATCCGAGAGCTTCTCTCTGAAAGGGTTTGAACATGCCAAAT 625
QY 482 TTGTGCTCTGTCATGACATATATCCACACCAAGAGACACTGACATTCGTTTTGAATAC 541
DB 626 TTGTGCTCTGTCATGACATATATCCACACCAAGAGACACTGACATTCGTTTTGAATAC 685
QY 542 TGACACACAGACTGGCCCACTATATGTCTCAGCATCCAGAGGGCTTCATCTCATATG 601
DB 686 TGACACACAGACTGGCCCACTATATGTCTCAGCATCCAGAGGGCTTCATCTCATATG 745
QY 602 TCAGACTTTTCAATGTTCAACTTTTGGGGGCTGGGCTGATCCACCAACCAAGGTC 661
DB 746 TCAGACTTTTCAATGTTCAACTTTTGGGGGCTGGGCTGATCCACCAACCAAGGTC 805
QY 662 TTACACAGGAGCTGAAACCTCAGAACTTACTCATCTGACCTGAGAGAGCTCAAACTGG 721
DB 806 TTACACAGGAGCTGAAACCTCAGAACTTACTCATCTGACCTGAGAGAGCTCAAACTGG 865
QY 722 CTGATTTTGTGCTTGGCCCGGCGCAAGTCCATTCACAGCAGACATACCTTTCAGAAAGTCG 781
DB 866 CTGATTTTGTGCTTGGCCCGGCGCAAGTCCATTCACAGCAGACATACCTTTCAGAAAGTCG 925
QY 782 TGACCTCTGTCATCCGGGCCCCCTGATGCTTGTGGGAGGCACTGAATATCTCTGACG 841
DB 926 TGACCTCTGTCATCCGGGCCCCCTGATGCTTGTGGGAGGCACTGAATATCTCTGACG 985
QY 842 TGACATATGAGGCTGACGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 901
DB 986 TGACATATGAGGCTGACGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 1045
QY 902 CTGGGGTTTTCACATCTCTTGAACAGCTGAGAGAAATCTGGAGGCTGTTGGAGTCCCTA 961
DB 1046 CTGGGGTTTTCACATCTCTTGAACAGCTGAGAGAAATCTGGAGGCTGTTGGAGTCCCTA 1105
QY 962 CAGAGGATCTTGGCCGGGAGTCTCAAGTACTTAACATTCAGAAATGCT 1015
DB 1106 CAGAGGATCTTGGCCGGGAGTCTCAAGTACTTAACATTCAGAAATGCT 1159
RESULT 12
ADM16422
ID ADM16422 standard; DNA; 1828 BP.
XX
XX ADM16422;
AC
XX
XX 17-JUN-2004 (first entry)
DT
XX
XX DNA encoding human kinase protein, seq id 1.
DE
XX
XX Cytostatic; cardiant; neuroprotective; nootropic; antinfertility;
XX vulnery; antidiabetic; kinase; cancer; heart disease;
XX Alzheimer's disease; infertility; wound; diabetes; neurological disease;
XX gene; de.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 87..1124
XX FT /*tag= a
XX FT /product= "kinase"
XX
XX WO2004024913-A1.
XX
XX 25-MAR-2004.
XX
XX 10-SEP-2003; 2003WO-JP011552.
XX
```

```
XX
PR 10-SEP-2003; 2002JP-00264345.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA (ZOEBS-) ZOEGENE CORP.
XX
XX Kondo J, Kawai K, Miyama N, Nakajima M, Isogai T, Sugiyama T;
PI Wakamatsu A, Irie R, Ishii S;
XX
XX MPI; 2004-270042/25.
DR P-PSDB; ADM16427.
XX
XX Proteins of human origin having kinase activity, useful for prevention
PT and treatment of kinase-associated diseases including cancer, heart
PT disease and Alzheimer's disease.
XX
XX Claim 4; SEQ ID NO 1; 105bp; Japanese.
XX
XX The invention relates to four proteins of human origin (I) having kinase
CC activity, and to proteins derived from these by addition, deletion and/or
CC substitution of one or more amino acid residues, and having similar
CC activity. The proteins and other aspects of the invention are useful for
CC the prevention and treatment of kinase-associated diseases including
CC cancer, heart disease, Alzheimer's disease, infertility, wounds, diabetes
CC and neurological diseases. The current sequence represents a human DNA
CC encoding a protein having kinase activity.
XX
XX Sequence 1828 BP; 496 A; 398 C; 450 G; 484 T; 0 U; 0 Other;
SQ
Query Match 66.7%; Score 873; DB 12; Length 1828;
Best Local Similarity 93.8%; Pred. No. 9.8e-269;
Matches 946; Conservative 0; Mismatches 0; Indels 63; Gaps 1;
QY 1 ATGGGTCAGAGCTGTGTCAGAACTGTAACGCTGATGAGCTGTACCATTTGTTCA 60
DB 87 ATGGGTCAGAGCTGTGTCAGAACTGTAACGCTGATGAGCTGTACCATTTGTTCA 146
QY 61 GAGGAGGCGGAGGACACACAGCTGTGAGAGAGTCAAGCTTGAGACCAAGAGGCTGCTTC 120
DB 147 GAGGAGGCGGAGGACACACAGCTGTGAGAGAGTCAAGCTTGAGACCAAGAGGCTGCTTC 206
QY 121 AAGCTTAACAGACTTAAGAAACATCATGTTCCATGACTTCTTCAACCCAGGGGACTT 180
DB 207 AAGCTTAACAGACTTAAGAAACATCATGTTCCATGACTTCTTCAACCCAGGGGACTT 266
QY 181 CAAGCTGCCCGTCCAGAGTTCAAGAGTAAAGGCCAGGAGTAAACAGTATGTTT 240
DB 267 CAAGCTGCCCGTCCAGAGTTCAAGAGTAAAGGCCAGGAGTAAACAGTATGTTT 326
QY 241 CAGGAAGAGGATCTGAGCGAGGGTTTTCAGTGGAGGAAGGCTCTTTTGGGGCAGCC 300
DB 327 CAGGAAGAGGATCTGAGCGAGGGTTTTCAGTGGAGGAAGGCTCTTTTGGGGCAGCC 386
QY 301 TCATCTTACTTGAACCTTGGAGAACTGGGTGAAGGCTCTTAAGCCACATTTAACAAGGG 360
DB 387 TCATCTTACTTGAACCTTGGAGAACTGGGTGAAGGCTCTTAAGCCACATTTAACAAGGG 446
QY 361 ATTAGAGAAATTAATGACAACTAGTGGCTTTAAAGTCAATCAGCATGAAATCAGAGAA 420
DB 447 ATTAGAGAAATTAATGACAACTAGTGGCTTTAAAGTCAATCAGCATGAAATCAGAGAA 506
QY 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGTAAGGGTTTGAACATGCCAAT 480
DB 507 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGTAAGGGTTTGAACATGCCAAT 566
QY 481 ATTGTGCTCTGTCATGACATATATCCACCAAGAGACATGACATTCGTTTTGAATAC 540
DB 567 ATTGTGCTCTGTCATGACATATATCCACCAAGAGACATGACATTCGTTTTGAATAC 626
QY 541 ATGACACAGAGCTGGCCAGATATATGTCAGCATCCAGAGGGCTTCATCTCATATAT 600
DB 627 ATG----- 629
```

QY 601 GTCAGACTTTTCACTTTTCACTTTTGGGGGCTGGGCTGACATCCAGCAACCAACGCTT 660
 DB 630 -----CTTTTCAATGTTTCAACTTTTGGGGGCTGGGCTGACATCCAGCAACCAACGCTT 683
 QY 661 CTTACAGAGGACCTGAAACCTCAGAACTTACTCATGATCATCCTGGAGAGCTCAAACTG 720
 DB 684 CTTACAGAGGACCTGAAACCTCAGAACTTACTCATGATCATCCTGGAGAGCTCAAACTG 743
 QY 721 GCTGATTTTGTGCTTGGCCCGGGGCAAGTCATTTCCAGCCAGCATATCTTTCAAGATC 780
 DB 744 GCTGATTTTGTGCTTGGCCCGGGGCAAGTCATTTCCAGCCAGCATATCTTTCAAGATC 803
 QY 781 GTGACCTCTGTGATCCGGCCCGCTGATGCTTGGAGGCACTGAATATTCCTGTGAG 840
 DB 804 GTGACCTCTGTGATCCGGCCCGCTGATGCTTGGAGGCACTGAATATTCCTGTGAG 863
 QY 841 CTTGACATATGAGGGTCAAGCTGCATCTTTATTTGAATGTTCCAGGGTCAACTTTGTTT 900
 DB 864 CTTGACATATGAGGGTCAAGCTGCATCTTTATTTGAATGTTCCAGGGTCAACTTTGTTT 923
 QY 901 CTTGGGGTTCCTCAACATCTTTGAACAGCTGGAGAAAATCTGGAGGTGCTGGAGTCCCT 960
 DB 924 CTTGGGGTTCCTCAACATCTTTGAACAGCTGGAGAAAATCTGGAGGTGCTGGAGTCCCT 983
 QY 961 ACAGAGGATCTTGGCCGGAGTCTCCAGCTACTCACTCAATCAG 1009
 DB 984 ACAGAGGATCTTGGCCGGAGTCTCCAGCTACTCACTCAATCAG 1032

RESULT 13

AB277160
 ID AB277160 standard, cDNA; 888 BP.

AC AB277160;
 DT 07-MAY-2003 (first entry)

XX Human protein kinase encoding cDNA SEQ ID NO:69.

XX Human; protein kinase; enzyme; antidiabetic; antiinflammatory;
 XX antidiabetic; antiparkinsonian; antimigraine; cardiant; cyostatic;
 XX immunosuppressive; vulnereary; gene therapy; COPD; asthma; migraine;
 XX chronic obstructive pulmonary disease; non-insulin dependent diabetes;
 XX Parkinson's disease; myocardial infarction; inflammatory bowel disease;
 XX autoimmune disorder; allograft rejection; graft versus host disease;
 XX cancer; leukemia; wound granulation; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..888

XX /tag= a

XX /partial

XX /product= "protein kinase"

XX /note= "no start or stop codons given"

XX WO2003000901-A2.

XX 03-JAN-2003.

XX 24-JUN-2002; 2002WC-IB002358.

XX 26-JUN-2001; 2001US-0301098P.

XX 06-NOV-2001; 2001US-0332870P.

XX (DECO-) DECODE GENETICS EHF.

XX Martinez RAM, Sigurdsson GT;

XX WPI; 2003-201429/19.

XX P-PsDB; ABP96082.

XX New protein kinase genes and polypeptides, useful for diagnosing diseases

PT associated with a protein kinase, or in gene therapy for treating e.g.
 PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
 PT or cancers.

XX Claim 1; Page 85; 258pp; English.

XX PS AB277126 to AB277165 encode the human protein kinases given in ABP96048
 CC to ABP96087. The protein kinases have antidiabetic, antiinflammatory,
 CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cyostatic,
 CC immunosuppressive and vulnereary activities, and can be used in gene
 CC therapy. A protein kinase therapeutic agent from the present invention,
 CC particularly a protein kinase gene agonist or antagonist, can be used for
 CC treating a disease or condition associated with a protein kinase in an
 CC individual. These diseases include chronic obstructive pulmonary diseases
 CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
 CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
 CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
 CC (e.g. leukemia) or wound granulation

XX Sequence 888 BP; 214 A; 234 C; 207 G; 233 T; 0 U; 0 Other;

XX Query Match 62.8%; Score 821; DB 10; Length 888;

XX Best Local Similarity 95.7%; Pred. No. 3.2e-252;

XX Matches 888; Conservative 0; Mismatches 30; Indels 9; Gaps 1;

QY 298 GCTCATCTTACTTGAAGCTGGAGAGGCTGTTAGCGACAGTTTACAG 357
 DB 1 GCTCATCTTACTTGAAGCTGGAGAGGCTGTTAGCGACAGTTTACAG 60
 QY 358 GGGATTAGAGAAATTAATGAGCACTAGTGGCTTTAAAGTCATGAGTGAATGAGAG 417
 DB 61 GGGATTAGAGAAATTAATGAGCACTAGTGGCTTTAAAGTCATGAGTGAATGAGAG 120
 QY 418 GAAGAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGGTTGAATATGCC 477
 DB 121 GAAGAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGGTTGAATATGCC 180
 QY 478 AATATTGGCTCTGTCATGATTAATCCAGCAAGAGAGCACTGACATTCGTTTGA 537
 DB 181 AATATTGGCTCTGTCATGATTAATCCAGCAAGAGAGCACTGACATTCGTTTGA 240
 QY 538 TACATGACACAGAGCTGGCCAGTATATGTCAGCATCCAGAGGCTTATCTCAT 597
 DB 241 TACATGAGTATGTTACCTGGCCAGTATATGTCAGCATCCAGAGGCTTATCTCAT 300
 QY 598 AATGTCAGCTTTTCACTGTTTCACTTTGGCGGGCTGGCTATCCACCAAGAC 657
 DB 301 AATGTCAGCTTTTCACTGTTTCACTTTGGCGGGCTGGCTATCCACCAAGAC 360
 QY 658 GTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATGATCATCCTGGAGAGCTCA 717
 DB 361 GTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATGATCATCCTGGAGAGCTCA 420
 QY 718 CTGGCTGATTTTGTCTTGGCCCGGGCAAGTCATTTCCAGCCAGCATATCTTTCAG 777
 DB 421 CTGGCTGATTTTGTCTTGGCCCGGGCAAGTCATTTCCAGCCAGCATATCTTTCAG 480
 QY 778 GTGTGACCTCTGTGTAACCGCCCTGTATGCTTTGCTGGAGCCATGAATATTCCT 837
 DB 481 GTGTGACCTCTGTGTAACCGCCCTGTATGCTTTGCTGGAGCCATGAATATTCCT 540
 QY 838 GAGCTGACATATGAGGGTGCAGAGCTGCATCTTATTTGAATGTTCCAGGGTCAACTTTG 897
 DB 541 GAGCTGACATATGAGGGTGCAGAGCTGCATCTTATTTGAATGTTCCAGGGTCAACTTTG 600
 QY 898 TTTCTGTGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTTC 957
 DB 601 TTTCTGTGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTTC 660
 QY 958 CTTACAGAGGATCTTGGCCGGAGTCTTCAAGCTTACTCACTCAATCAATCAAGATGTT 1017
 DB 661 CTTACAGAGGATCTTGGCCGGAGTCTTCAAGCTTACTCACTCAATCAATCAAGATGTT 720

QY 1018 CCAGTCCCTGAGGCTGGAAGCCTTATGTTGCTGGAAGAGGCTGGGAGGGTTCCTGAA 1077
 DB 721 AAAAACTATATTTTCTTCC-----ACTTTTCCAGGCTGGGAGGGTTCCTGAA 771
 QY 1078 GGTGAAGACCTGGCCCTCCAGATGCTPAAAGGCTTCCAGAGACCGCTTCGCCGCA 1137
 DB 772 GGTGAAGACCTGGCCCTCCAGATGCTPAAAGGCTTCCAGAGACCGCTTCGCCGCA 831
 QY 1138 GAAGCATTTGTTATATTTTACAGCGCCCTGCACTGATGACAGCTTCT 1194
 DB 832 GAAGCATTTGTTATATTTTACAGCGCCCTGCACTGATGACAGCTTCT 888
 RESULT 14
 AAF44668
 ID AAF44668 standard; cDNA; 1077 BP.
 AC AAF44668;
 XX
 XX 27-MAR-2001 (first entry)
 DT
 XX Novel protein kinase cDNA, SEQ ID NO: 48.
 DE
 XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KM immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KM dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KM immune disorder; cardiovascular disease; neurodegenerative disease;
 KM cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KM inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX Homo sapiens.
 OS
 XX WO200073469-A2.
 PN
 PD 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000MO-US014842.
 PF
 XX 28-MAY-1999; 99US-0136503P.
 PR
 XX (SUGB-) SUGEN INC.
 PA
 XX Plowman GD, Martinez R, Whyte D, Suderanan S;
 PI WPI; 2001-032161/04.
 DR P-PSDB; AAB65641.
 DR
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers.
 PT
 XX Example 1; Fig 2; 310pp; English.
 PS
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
 CC stress related disorders, chronic inflammatory bowel disease, chronic
 CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
 CC psoriasis, thinitis, autoimmunity, diabetes, cancers and reproductive
 CC disorders
 XX
 SO Sequence 1077 BP; 269 A; 257 C; 248 G; 303 T; 0 U; 0 Other;

Query Match 53.7%; Score 702.2; DB 4; Length 1077;
 Best Local Similarity 99.6%; Pred. No. 4.9e-214;
 Matches 704; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 602 TCAGACTTTTCATGTTTCAACTTTTGGCGGGCTGGCGTATCCACCAACAGTTTC 661
 DB 95 TCAGACTTTTCATGTTTCAACTTTTGGCGGGCTGGCGTATCCACCAACAGTTTC 154
 QY 662 TTCACAGGAGACCTGAAACCTGAAACTTACTATCATGTCACCTGGGAGAGCTCAACCTG 721
 DB 155 TTCACAGGAGACCTGAAACCTGAAACTTACTATCATGTCACCTGGGAGAGCTCAACCTG 214
 QY 722 CTGATTTTGGTCTTGGCGGGGCAAGTCCATTCGCCAGCAGCATACTCTTCAGAACTG 781
 DB 215 CTGATTTTGGTCTTGGCGGGGCAAGTCCATTCGCCAGCAGCATACTCTTCAGAACTG 274
 QY 782 TGACCTCTGTATCCGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGTAGC 841
 DB 275 TGACCTCTGTATCCGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGTAGC 334
 QY 842 TGGACATATGGGCTGAGGCTGATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTTC 901
 DB 335 TGGACATATGGGCTGAGGCTGATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTTC 394
 QY 902 CTGGGGTTTCAACATCCTTGAACAGCTGGAGAAATCTGGAGGCTGGAGTCCCTA 961
 DB 395 CTGGGGTTTCAACATCCTTGAACAGCTGGAGAAATCTGGAGGCTGGAGTCCCTA 454
 QY 962 CAGAGATATCTTGGCGGGAGTCTCAAGCTTACTACTCAATCCAGAAATGTTCCAC 1021
 DB 455 CAGAGATATCTTGGCGGGAGTCTCAAGCTTACTACTCAATCCAGAAATGTTCCAC 514
 QY 1022 TGCCTACGCTCTGAAGCCTTCAATGTTTCTGGAACAGCTGGGAGGTTCTCTGAAGCTG 1081
 DB 515 TGCCTACGCTCTGAAGCCTTCAATGTTTCTGGAACAGCTGGGAGGTTCTCTGAAGCTG 574
 QY 1082 AAGACCTGGCTCCAGATGCTPAAAGGCTTCCAGAGACCGCGTCTCGGCCAGGA 1141
 DB 575 AAGACCTGGCTCCAGATGCTPAAAGGCTTCCAGAGACCGCGTCTCGGCCAGGA 634
 QY 1142 CACTTGTTCATGTTTATTTTACAGCGCCCTGCACTCTGAGCTGATGAGT 1201
 DB 635 CACTTGTTCATGTTTATTTTACAGCGCCCTGCACTCTGAGCTGATGAGT 694
 QY 1202 AGCTTTGTTTACAGTTTACAGAGTGAAGCTTAAAGCCAGAAATGTTGACCTTTGGCCT 1261
 DB 695 AGCTTTGTTTACAGTTTACAGAGTGAAGCTTAAAGCCAGAAATGTTGACCTTTGGCCT 754
 QY 1262 CCTACAGAAAGTCAACCCAGGCCAGTTTAAAGAAATGCTGATGA 1308
 DB 755 CCTACAGAAAGTCAACCCAGGCCAGTTTAAAGAAATGCTGATGA 801
 RESULT 15
 ADI29366
 ID ADI29366 standard; cDNA; 1077 BP.
 XX
 XX ADI29366;
 AC
 XX 22-APR-2004 (first entry)
 DT
 XX Human MARK3-associated cDNA #36.
 DE
 XX Human; ss; antisense gene therapy; MARK3;
 KM MAP/microtubule affinity-regulating kinase 3; cancer;
 KM Alzheimer's disease; neurodegenerative disorder;
 KM hyperproliferative disorder; cytostatic.
 XX Homo sapiens.
 OS
 XX US2003232771-A1.
 PN
 XX 18-DEC-2003.

XX 17-JUN-2002; 2002US-00174319.
XX 17-JUN-2002; 2002US-00174319.
XX (ISIS-) ISIS PHARM INC.
XX Ward DT, Freier SM, Dobie KW;
XX WPI; 2004-052188/05.
XX P-PSDB; ADI29248.
XX
XX New antisense compound targeted to a nucleic acid molecule encoding
XX microtubule-affinity-regulating kinases (MARK3), useful for modulating
XX expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
XX Disclosure; Fig 2; 233p; English.
XX
XX The invention relates to a compound comprising a sequence comprising 8-80
XX base pairs (bp) targeted to a nucleic acid encoding MARK3
XX (MAP/microtubule affinity-regulating kinase 3), that specifically
XX hybridizes with the nucleic acid encoding MARK3 and inhibits expression
XX of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
XX composition comprising the compound and a carrier or diluent, inhibiting
XX the expression of MARK3 in cells or tissues, treating an animal having or
XX suspected of having a disease or condition associated with MARK3 and
XX screening for an antisense compound. The antisense oligonucleotide is
XX useful for preparing a composition for treating hyperproliferative
XX disorder, particularly cancer and neurodegenerative diseases e.g.
XX Alzheimer's disease. The present sequence is a MARK3 associated cDNA
XX included in the figures but not mentioned anywhere else in the
XX specification.
XX
XX Sequence 1077 BP; 269 A; 257 C; 248 G; 303 T; 0 U; 0 Other;

Query Match 53.7%; Score 702.2; DB 12; Length 1077;
Best Local Similarity 99.6%; Pred. No. 4.9e-214;
Matches 704; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 602 TCAGACTTTTCATGTTTCACTTTTGCGGGGCTGGGGTACATCCACCAACACGTTT 661
DB 95 TCCAGCTTTTCATGTTTCACTTTTGCGGGGCTGGGGTACATCCACCAACACGTTT 154
QY 662 TTCACAGGAGACCTGAAACCTCAACTTACTCATGTCACCTGGGAGAGCTCAAACTGG 721
DB 155 TTACACAGGAGACCTGAAACCTCAACTTACTCATGTCACCTGGGAGAGCTCAAACTGG 214
QY 722 CTGATTTTGGTCTTGCCCGGGCCAAAGTCCATTTCCAGCCAGACATCTTTGAGAAGTCG 781
DB 215 CTGATTTTGGTCTTGCCCGGGCCAAAGTCCATTTCCAGCCAGACATCTTTGAGAAGTCG 274
QY 782 TGACCTCTGGTACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTTGAGC 841
DB 275 TGACCTCTGGTACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTTGAGC 334
QY 842 TGGACATATGCGGGTGAAGGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 901
DB 335 TGGACATATGCGGGTGAAGGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 394
QY 902 CTGGGGTTTCCAAATCTTTGAAACAGCTGGAGAAAATCTGGAGAGTCTGGAGTCCCTA 961
DB 395 CTGGGGTTTCCAAATCTTTGAAACAGCTGGAGAAAATCTGGAGAGTCTGGAGTCCCTA 454
QY 962 CAGAGATATCTGGCGGGAGTCTCCAAAGTACTTAATCAATCCAGAAATGTTCCAC 1021
DB 455 CAGAGATATCTGGCGGGAGTCTCCAAAGTACTTAATCAATCCAGAAATGTTCCAC 514
QY 1022 TGGCTAAGCCTGAAAGCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAGCTG 1081
DB 515 TGGCTAAGCCTGAAAGCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAGCTG 574
QY 1082 AAGACCTGGCCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGAAG 1141

DB 575 AAGACCTGGCCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGAAG 634
QY 1142 CACTTGTTCATGATATTTTCAAGCGCCCTGCAATCTCACTGATGATGAG 1201
DB 635 CACTTGTTCATGATATTTTCAAGCGCCCTGCAATCTCACTGATGATGAG 694
QY 1202 AGCTTTGTTTACAGTTTCAGAGTGAAGGCTAAAGCAGAAATGTTGACCTTTGGGCT 1261
DB 695 AGCTTTGTTTACAGTTTCAGAGTGAAGGCTAAAGCAGAAATGTTGACCTTTGGGCT 754
QY 1262 CCTACCAAGAAAGTCAACCAAGCCCAAGTTTACCAATGCTGTGA 1308
DB 755 CCTACCAAGAAAGTCAACCAAGCCCAAGTTTACCAATGCTGTGA 801

Search completed: December 26, 2004, 18:49:05
Job time : 382.27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:39:26 ; Search time 3570.15 Seconds
(without alignments)
17325.567 Million cell updates/sec

Title: US-10-786-065-4

Perfect score: 1 atgggtcgaagcgcgtgc.....agttacgaacatgcgtgta 1308

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenBank:*

1: gb_da:*\n2: gb_hcg:*\n3: gb_in:*\n4: gb_om:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_srb:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1308	100.0	1308	6	AR265352	Sequence
2	1308	100.0	1308	6	AX571876	Sequence
3	1308	100.0	1308	6	AX772798	Sequence
4	1308	100.0	1584	6	AX772798	Sequence
5	1308	100.0	1628	6	AX772798	Sequence
6	1243.6	95.1	3675	9	HSN807539	Sequence
7	1200	91.7	2305	9	BC038807	Homo sapi
8	1185.4	90.6	1534	6	AX772791	Sequence
9	1185.4	90.6	1534	6	AB053308	Homo sapi
10	1185.4	90.6	1534	6	AB053308	Homo sapi
11	1009	77.1	2203	6	AR265350	Sequence
12	1009	77.1	2203	6	AR265350	Sequence
13	887.6	67.9	1376	6	AX644241	Sequence
14	873	66.7	1828	9	AK131512	Homo sapi
15	821	62.8	888	6	AX803407	Sequence
16	702.2	53.7	1077	6	AX056403	Sequence
17	481	36.8	481	6	AR175616	Sequence
18	481	36.8	481	6	AR236733	Sequence
19	439.6	33.6	4529	6	CQ777528	Sequence

20	439.6	33.6	4529	10	AF033655	Mus muscu
21	438	33.5	2467	10	MMU62391	Mus muscu
22	438	33.5	4732	10	AK129226	Mus muscu
23	438	33.5	5127	10	BC068134	Mus muscu
24	436.2	33.3	4957	6	CQ776662	Sequence
25	436.2	33.3	4957	6	AX925598	Sequence
26	436.2	33.3	4957	9	AB020641	Homo sapi
27	434.6	33.2	2043	9	AF119833	Homo sapi
28	432.2	33.0	748	6	CQ715966	Sequence
29	404.2	30.9	2487	5	BC075148	Xenopus la
30	401.6	30.7	1220	5	BX929373	Gallus ga
31	399.2	30.5	1314	6	CQ717496	Sequence
32	378.6	28.9	1196	5	BX932969	Gallus ga
33	339.4	25.9	2198	3	AY113367	Drosophi
34	339.4	25.9	4789	3	BT010064	Drosophi
35	337.8	25.8	2641	3	AF152403	Drosophi
36	337.8	25.8	2755	3	AF152404	Drosophi
37	337.8	25.8	2933	3	AF152402	Drosophi
38	337.8	25.8	2942	3	AF152406	Drosophi
39	337.8	25.8	2972	3	AF152401	Drosophi
40	337.8	25.8	3065	3	AF152400	Drosophi
41	337.8	25.8	3346	3	AF152405	Drosophi
42	337.8	25.8	3617	3	AF152399	Drosophi
43	337.8	25.8	3656	3	AF152398	Drosophi
44	334.6	25.6	3526	3	DMPF7AIRE	X99512 D.melanog
45	329.2	25.2	2703	6	CQ600717	Sequence

ALIGNMENTS

RESULT 1
LOCUS AR265352 1308 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 4 from patent US 6492154.
ACCESSION AR265352
VERSION AR265352.1 GI:29693855
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1308)
AUTHORS Yan, C., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, and uses thereof
human kinase proteins, and uses thereof
JOURNAL Patent: US 6492154-A 4 10-DEC-2002;
FEATURES
source 1..1308
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1308; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGAGCTGTACCATTTGCA 60
DB 1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGAGCTGTACCATTTGCA 60
QY 61 GAGGGAGGCGAGGACACACAGCTGTGTGAGAGATCAGCTGTGAGAGAGAGCTGTG 120
DB 61 GAGGGAGGCGAGGACACACAGCTGTGTGAGAGATCAGCTGTGAGAGAGAGCTGTG 120
QY 121 AAGCTTAAGAGCTTAAGAGATCATGTTTCATGACTTCACTTCACTTCACTTCACTT 180
DB 121 AAGCTTAAGAGCTTAAGAGATCATGTTTCATGACTTCACTTCACTTCACTTCACTT 180
QY 181 CAAAGTCCCGTCCCGAGAGTTCAAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 CAAAGTCCCGTCCCGAGAGTTCAAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 CAGGAAGAGAGATCTGAGGAGGAGTTTTCAGTGTGAGAGAGAGAGCTCCCTTTTGGGGAGGCC 300
DB 241 CAGGAAGAGAGATCTGAGGAGGAGTTTTCAGTGTGAGAGAGAGAGCTCCCTTTTGGGGAGGCC 300

Db	241	CAGGAAGAAGATCTGAGGCAAGGTTTTACGTGAGAAAGAGCTTCCCTTTGGGGCAACC	300
Qy	301	TCATCTTACTTGAACCTTGAGAGAAGCTGGGTGAAGGCTCTTATGCGACAGTTTACAAAGGG	360
Db	301	TCATCTTACTTGAACCTTGAGAGAAGCTGGGTGAAGGCTCTTATGCGACAGTTTACAAAGGGG	360
Qy	361	ATTAGCAGAAATAAATGACAACCTAGTGGCTTTAAAAGTCATCAGAGTAAATGCAGAGAA	420
Db	361	ATTAGCAGAAATAAATGACAACCTAGTGGCTTTAAAAGTCATCAGAGTAAATGCAGAGAA	420
Qy	421	GGAGTCCCATTTTACAGCTTATCCGAGAAGCTTCTCTCTGAAAGGTTGAAAATGCCAAT	480
Db	421	GGAGTCCCATTTTACAGCTTATCCGAGAAGCTTCTCTCTGAAAGGTTGAAAACATCCCAAT	480
Qy	481	ATTGAGCTCCGCAATGACATTAATCCACAACAAAGAGACACTGACATTCGTGTTTTGAATAC	540
Db	481	ATTGAGCTCCGCAATGACATTAATCCACAACAAAGAGACACTGACATTCGTGTTTTGAATAC	540
Qy	541	ATGCAACACAGACTTGAGCCCAAGTATATGTCAGACATCCAGAGAGGCTTCATCTCATAT	600
Db	541	ATGCAACACAGACTTGAGCCCAAGTATATGTCAGACATCCAGAGAGGCTTCATCTCATAT	600
Qy	601	GTGACACTTTTCAGTGTTCACATTTTGGGGGGCCCTGGGGTACATCCACACCAACAGCTT	660
Db	601	GTGACACTTTTCAGTGTTCACATTTTGGGGGGCCCTGGGGTACATCCACACCAACAGCTT	660
Qy	661	CTTCACAGGAGACTGAAAACCTCAGAACCTTACTCATCATCTGAGAGAGGCTCAAACTG	720
Db	661	CTTCACAGGAGACTGAAAACCTCAGAACCTTACTCATCATCTGAGAGAGGCTCAAACTG	720
Qy	721	GCTGATTTTGGTCTTTGGCCGGGGCCAAATCCATTTCCAGCCAGACATCTTTCCAGAAGTC	780
Db	721	GCTGATTTTGGTCTTTGGCCGGGGCCAAATCCATTTCCAGCCAGACATCTTTCCAGAAGTC	780
Qy	781	GTGACCCCTCTGTATCCGGGCCCTCTGATCTTTTGGGGAGGCACTGAATATTTCCCTGAG	840
Db	781	GTGACCCCTCTGTATCCGGGCCCTCTGATCTTTTGGGGAGGCACTGAATATTTCCCTGAG	840
Qy	841	CTGACATATGAGGGGTGAGGCTGCATCTTTATGGAATGTTCCAGAGGTCAACTTTGT	900
Db	841	CTGACATATGAGGGGTGAGGCTGCATCTTTATGGAATGTTCCAGAGGTCAACTTTGT	900
Qy	901	CCTGGGGTTTCCACATCTCTGAAACAGCTGAGAAAATCTGGAGAGTCTGGAGTCCCT	960
Db	901	CCTGGGGTTTCCACATCTCTGAAACAGCTGAGAAAATCTGGAGAGTCTGGAGTCCCT	960
Qy	961	ACAGAGGATACCTTGGCCGGAGGTCTCAAGCTACTTAATCTCAATCCAGAAATGTTCCCA	1020
Db	961	ACAGAGGATACCTTGGCCGGAGGTCTCAAGCTACTTAATCTCAATCCAGAAATGTTCCCA	1020
Qy	1021	CTGGCTTACGCTCTGAAAGCTTCATGTTGTCTGGAACAGGCTGGGCAAGGTTCTGAACT	1080
Db	1021	CTGGCTTACGCTCTGAAAGCTTCATGTTGTCTGGAACAGGCTGGGCAAGGTTCTGAACT	1080
Qy	1081	GAAACACCTGAGCCTCCAGATGCTAAAGAGCTTCCACAGAGACCGGCTTCCGCCAGAAA	1140
Db	1081	GAAACACCTGAGCCTCCAGATGCTAAAGAGCTTCCACAGAGACCGGCTTCCGCCAGAAA	1140
Qy	1141	GCACCTGTTCAATGATTAATTTTCAAGCGCCCTGCAATCTCAGCTGACACCTTCTGATGAG	1200
Db	1141	GCACCTGTTCAATGATTAATTTTCAAGCGCCCTGCAATCTCAGCTGACACCTTCTGATGAG	1200
Qy	1201	GAGCTTTTGTATTAAGTTTACAGAGTGAAGGCTTAAGCCAGAAATGTTGACCTTTTGACC	1260
Db	1201	GAGCTTTTGTATTAAGTTTACAGAGTGAAGGCTTAAGCCAGAAATGTTGACCTTTTGACC	1260
Qy	1261	TCCTTACCAAGAAAGTCAACACCCAGGCCAGTTTACCAATGCTGGTGA 1308	
Db	1261	TCCTTACCAAGAAAGTCAACACCCAGGCCAGTTTACCAATGCTGGTGA 1308	

AX571876	1308 bp	DNA	linear	PAT 29-NOV-2002		
LOCUS	AX571876	Sequence 4 from Patent WO02061060.				
DEFINITION	AX571876					
ACCESSION	AX571876.1	GI:26004001				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
ORIGIN						
Query Match	100.0%;	Score 1308;	DB 6;	Length 1308;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1308;	Conservative	0;	Mismatches	0;		
			Indels	0;		
			Gaps	0;		
QY	1	ATGGGTCAAGAGCTGTGTGC	AAAGACGTGACAGCTGGATG	CAGCTGCTACATTGTTCA	60	
DB	1	ATGGGTCAAGAGCTGTGTGC	AAAGACGTGACAGCTGGATG	CAGCTGCTACATTGTTCA	60	
QY	61	GAGGGAGCGCAGGACAC	ACAGCTGTGAGGAGT	CAGCTTGACACGAGGCTGC	120	
DB	61	GAGGGAGCGCAGGACAC	ACAGCTGTGAGGAGT	CAGCTTGACACGAGGCTGC	120	
QY	121	AAGCTTACAGACCTTAA	AAAGACATCATGTTCCAT	TGACCTTCAACCCAGGGGACTT	180	
DB	121	AAGCTTACAGACCTTAA	AAAGACATCATGTTCCAT	TGACCTTCAACCCAGGGGACTT	180	
QY	181	CAAGCTCCCGTCCGAC	GAAGTTCAAGAGTAAAG	GGCCACGAGTAAACATGATTTGTTT	240	
DB	181	CAAGCTCCCGTCCGAC	GAAGTTCAAGAGTAAAG	GGCCACGAGTAAACATGATTTGTTT	240	
QY	241	CAGGAAGAGATCTGAG	GCAGGGTTTTCAGTGGAG	GAAGACCTTCCCTTTGGGGCAGCC	300	
DB	241	CAGGAAGAGATCTGAG	GCAGGGTTTTCAGTGGAG	GAAGACCTTCCCTTTGGGGCAGCC	300	
QY	301	TCATCTTACTTGAAC	TCTGGAGAACGCTCTT	ATGCGACAGTTTAAAGGGG	360	
DB	301	TCATCTTACTTGAAC	TCTGGAGAACGCTCTT	ATGCGACAGTTTAAAGGGG	360	
QY	361	ATTGACGAATTAAT	TGACCACTATGTGGCTTT	AAAAGTCATCAGATGAATG	CAGAGAA	420
DB	361	ATTGACGAATTAAT	TGACCACTATGTGGCTTT	AAAAGTCATCAGATGAATG	CAGAGAA	420
QY	421	GGAGTCCCATTTA	CAGCTATCCGAGAGCTT	CTCTCCGGAAGGGTTT	GAACATGCCCAAT	480
DB	421	GGAGTCCCATTTA	CAGCTATCCGAGAGCTT	CTCTCCGGAAGGGTTT	GAACATGCCCAAT	480
QY	481	ATTGTGCTCTG	CATGACATATCCAC	CAAGAGACACTGACATTT	CGTTTTTGAATAC	540
DB	481	ATTGTGCTCTG	CATGACATATCCAC	CAAGAGACACTGACATTT	CGTTTTTGAATAC	540
QY	541	ATGACACACA	CCCGGCCAGATTAAT	GTCTCAGATCCAGAGGGCTT	CAATCCCATAT	600
DB	541	ATGACACACA	CCCGGCCAGATTAAT	GTCTCAGATCCAGAGGGCTT	CAATCCCATAT	600
QY	601	GTCAGACTTTT	CAATGTTTCAACTTTT	CGGGGGCTGCGCTGAT	CAATCCACAACAGGTT	660
DB	601	GTCAGACTTTT	CAATGTTTCAACTTTT	CGGGGGCTGCGCTGAT	CAATCCACAACAGGTT	660
QY	661	CTTCACAGGAG	CTGAAACCTCAG	ACTTACTCATCATG	TCACTTGGGAGAGCT	720
DB	661	CTTCACAGGAG	CTGAAACCTCAG	ACTTACTCATCATG	TCACTTGGGAGAGCT	720

Db 661 CTTACAGGAGACCTGAAACCTCAGAACTTATCATCATGACCTGGAGAGCTCAAACTG 720
Qy 721 GCTGATTTTGTCTTGGCCCGGGCCAAAGTCATTTCCAGCCAGACATATCTTTCAAGATC 780
Db 721 GCTGATTTTGTCTTGGCCCGGGCCAAAGTCATTTCCAGCCAGACATATCTTTCAAGATC 780
Qy 781 GTGACCTCTGTGATCCGGCCCCCTGATGCTTTTGTGGAGCCATGAAATTTCTCTGAG 840
Db 781 GTGACCTCTGTGATCCGGCCCCCTGATGCTTTTGTGGAGCCATGAAATTTCTCTGAG 840
Qy 841 CTGGAATATGAGGGTGCAGGCTGCATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTT 900
Db 841 CTGGAATATGAGGGTGCAGGCTGCATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTT 900
Qy 901 CTTGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Db 901 CTTGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Qy 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAATCAATCCAGATGTTTCCCA 1020
Db 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAATCAATCCAGATGTTTCCCA 1020
Qy 1021 CTGGCTACGCTCGAAGCCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Db 1021 CTGGCTACGCTCGAAGCCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Qy 1081 GAAGACCTGGCCTCCAGATGCTTAAAGGCTTTCAGAGACCCGCTCTCCGCCAGAGAA 1140
Db 1081 GAAGACCTGGCCTCCAGATGCTTAAAGGCTTTCAGAGACCCGCTCTCCGCCAGAGAA 1140
Qy 1141 GCACTTTGTTCAATATTTTCAAGCCCTGCTCCATCTCAAGCTGTACAGCTTCTGTATGAG 1200
Db 1141 GCACTTTGTTCAATATTTTCAAGCCCTGCTCCATCTCAAGCTGTACAGCTTCTGTATGAG 1200
Qy 1201 GAGCTTTGTTTACATTTTCAAGAGTGAAGCTTAAAGCCAGAAATGTTGATCTTTTGACC 1260
Db 1201 GAGCTTTGTTTACATTTTCAAGAGTGAAGCTTAAAGCCAGAAATGTTGATCTTTTGACC 1260
Qy 1261 TCTTACAGAAAGGTGACCAACCCAGCTTATAGCAATGCTGTGTA 1308
Db 1261 TCTTACAGAAAGGTGACCAACCCAGCTTATAGCAATGCTGTGTA 1308
RESULT 3
AX772798 1308 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 8 from Patent WO03046167.
DEFINITION AX772798
ACCESSION AX772798
VERSION AX772798.1 GI:32485025
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Koehler, R.H.
Regulation of human serine/threonine protein kinase
Patent: WO 03046167-A 8 05-JUN-2003;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
source
1. .1308
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="serine/threonine protein kinase"
ORIGIN
Query Match 100.0%; Score 1308; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGACGCTGTACCAATTTGTCA 60
|||||
Db 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGACGCTGTACCAATTTGTCA 60

Db 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGACGCTGTACCAATTTGTCA 60
Qy 61 GAGGAGGGGAGGACACACAGCTGTGAGAGTCAAGCTTGAAGCAACGAGGCTGCGTTC 120
Db 61 GAGGAGGGGAGGACACACAGCTGTGAGAGTCAAGCTTGAAGCAACGAGGCTGCGTTC 120
Qy 121 AAGCTTACAGACCTTAAAGAAACATCATGTTCCATTCATTTCAACCCCGAGGACCT 180
Db 121 AAGCTTACAGACCTTAAAGAAACATCATGTTCCATTCATTTCAACCCCGAGGACCT 180
Qy 181 CAAGCTGCCCGTGGCCAGAAAGTTCAAGAGTAAAGGCGACGAGAGTAAAGTATTTT 240
Db 181 CAAGCTGCCCGTGGCCAGAAAGTTCAAGAGTAAAGGCGACGAGAGTAAAGTATTTT 240
Qy 241 CAGGAAGAGGATCTGAGGAGGGTTTCAAGTGAAGAGAGGCTTCTTTTGGGAGGCT 300
Db 241 CAGGAAGAGGATCTGAGGAGGGTTTCAAGTGAAGAGAGGCTTCTTTTGGGAGGCT 300
Qy 301 TCATCTTACTTGAATCTGGAAGAGCTGGTGAAGGCTTTATGCCACATTTTCAAGGGG 360
Db 301 TCATCTTACTTGAATCTGGAAGAGCTGGTGAAGGCTTTATGCCACATTTTCAAGGGG 360
Qy 361 ATTAGAGAAATTAATGGAACAATGAGCTTTTAAAGTATGATGATGATGATGATGATG 420
Db 361 ATTAGAGAAATTAATGGAACAATGAGCTTTTAAAGTATGATGATGATGATGATGATG 420
Qy 421 GAGTCCCAATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAACATGCAAT 480
Db 421 GAGTCCCAATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAACATGCAAT 480
Qy 481 ATTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 ATTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 ATGCAACAGACCTGGCCAGATATATGCTCAGATCCAGAGGAGGCTTATCTCTAAT 600
Db 541 ATGCAACAGACCTGGCCAGATATATGCTCAGATCCAGAGGAGGCTTATCTCTAAT 600
Qy 601 GTGAGCTTTTATGTTTCACTTTTTCGAGGCTGAGCTGATCAGATCCAGCAACAGTT 660
Db 601 GTGAGCTTTTATGTTTCACTTTTTCGAGGCTGAGCTGATCAGATCCAGCAACAGTT 660
Qy 661 CTTACAGGAGACCTGAAACCTCAGAACTTATCTATCATGCTACCTGGGAGAGCTCAAC 720
Db 661 CTTACAGGAGACCTGAAACCTCAGAACTTATCTATCATGCTACCTGGGAGAGCTCAAC 720
Qy 721 GCTGATTTTGTCTTGGCCCGGGCCAAAGTCAATTTCCAGCAGACATCTTCCAGAAATC 780
Db 721 GCTGATTTTGTCTTGGCCCGGGCCAAAGTCAATTTCCAGCAGACATCTTCCAGAAATC 780
Qy 781 GTGACCTCTGTGATCCGCCCCCTGATGCTTGTGGAGCCAGCTGAATATTTCTGTAG 840
Db 781 GTGACCTCTGTGATCCGCCCCCTGATGCTTGTGGAGCCAGCTGAATATTTCTGTAG 840
Qy 841 CTGGAATATGAGGGTGCAGGCTGCATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTT 900
Db 841 CTGGAATATGAGGGTGCAGGCTGCATCTTATTTGAAATGTTCCAGGGTCAACCTTTGTTT 900
Qy 901 CTTGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Db 901 CTTGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Qy 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAATCAATCCAGATGTTTCCCA 1020
Db 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAATCAATCCAGATGTTTCCCA 1020
Qy 1021 CTGGCTACGCTCGAAGCCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Db 1021 CTGGCTACGCTCGAAGCCTTCAATGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Qy 1081 GAAGACCTGGCCTCCAGATGCTTAAAGGCTTTCAGAGACCCGCTCTCCGCCAGAGAA 1140
Db 1081 GAAGACCTGGCCTCCAGATGCTTAAAGGCTTTCAGAGACCCGCTCTCCGCCAGAGAA 1140

QY 1141 GCACCTTGTATGATTTATTTACAGGCGCTTCCCATCTTCAAGCTGTACCAAGCTTCTCTGATGAG 1200
Db 1141 GCACCTTGTATGATTTATTTACAGGCGCTTCCCATCTTCAAGCTGTACCAAGCTTCTCTGATGAG 1200
QY 1201 GAGCTTTTGTTCAGATTTCAGAGAGTGAAGGCTAAAGCCAGAAATGTGTACCTTTTGGCC 1260
Db 1201 GAGCTTTTGTTCAGATTTCAGAGAGTGAAGGCTAAAGCCAGAAATGTGTACCTTTTGGCC 1260
QY 1261 TCCCTACAGAAAGGTCCACCAAGCCCAAGTTAGCAAAATGCTGTGA 1308
Db 1261 TCCCTACAGAAAGGTCCACCAAGCCCAAGTTAGCAAAATGCTGTGA 1308

RESULT 4
AX772794 1584 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 4 from Patent WO03046167.
DEFINITION AX772794
ACCESSION AX772794
VERSION AX772794.1 GI:32485023
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase
PATENT: WO 03046167-A 4 05-JUN-2003;
JOURNAL Bayer Aktiengesellschaft (DE)
LOCATION/Qualifiers
1. 1584
FEATURES
source /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="serine/threonine protein kinase"

ORIGIN
Query Match 100.0%; Score 1308; DB 6; Length 1584;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 60
Db 1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 60
QY 61 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGTAGACCAAGAGGCTGCGTTC 120
Db 61 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGTAGACCAAGAGGCTGCGTTC 120
QY 121 AAGCTTAAGAGCTTAAGAAAGATCATGTTCCATGACTTTCATTTACCCCAAGGAGACTT 180
Db 121 AAGCTTAAGAGCTTAAGAAAGATCATGTTCCATGACTTTCATTTACCCCAAGGAGACTT 180
QY 181 CAAGTCCCGTCCGAGAAAGTCAAGAGTAAAGGCGACGAGTAACAGTATGTTTT 240
Db 181 CAAGTCCCGTCCGAGAAAGTCAAGAGTAAAGGCGACGAGTAACAGTATGTTTT 240
QY 241 CAGGAAGAGATCTGAGGAGGCTTTCAATGAGAGAAAGCCTCTTTTGGGCGAGCC 300
Db 241 CAGGAAGAGATCTGAGGAGGCTTTCAATGAGAGAAAGCCTCTTTTGGGCGAGCC 300
QY 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTATGCGACAGTTTACAAAGGG 360
Db 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTATGCGACAGTTTACAAAGGG 360
QY 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTCACTGCAATGATGAGGAGAA 420
Db 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTCACTGCAATGATGAGGAGAA 420
QY 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTCTGAAGGTTTGAACAATGCCAAT 480
Db 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTCTGAAGGTTTGAACAATGCCAAT 480

QY 481 ATTGTCCTCTGATGACATAATCCACACCAAGAGACCTGACATTCGTTTTGATAC 540
Db 481 ATTGTCCTCTGATGACATAATCCACACCAAGAGACCTGACATTCGTTTTGATAC 540
QY 541 ATGACACAGACCTGGCCAGATATATGTCTGAGCATCCGAGAGGGCTTATCCTCTAAT 600
Db 541 ATGACACAGACCTGGCCAGATATATGTCTGAGCATCCGAGAGGGCTTATCCTCTAAT 600
QY 601 GTGAGCTTTTCAATGTTTCACTTTTGCAGGCGCTGGCTGATATCCACCAACACGTT 660
Db 601 GTGAGCTTTTCAATGTTTCACTTTTGCAGGCGCTGGCTGATATCCACCAACACGTT 660
QY 661 CTTTCACAGGAGCTGAAAACCTCAGAACTTACTATCATGCTACCTGGAGAGCTCAACTG 720
Db 661 CTTTCACAGGAGCTGAAAACCTCAGAACTTACTATCATGCTACCTGGAGAGCTCAACTG 720
QY 721 GCTGATTTTGTCTTGCCCGGGGCAAGTCCATTCCAGCAGACATPACTTCAGAAATC 780
Db 721 GCTGATTTTGTCTTGCCCGGGGCAAGTCCATTCCAGCAGACATPACTTCAGAAATC 780
QY 781 GTGACCTCTGTGTACCGGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGAG 840
Db 781 GTGACCTCTGTGTACCGGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGAG 840
QY 841 CTGACATATGAGGCTGACAGGCTGCATCTTTATTTGAATGTTCCAGGCTCAACTTTGTT 900
Db 841 CTGACATATGAGGCTGACAGGCTGCATCTTTATTTGAATGTTCCAGGCTCAACTTTGTT 900
QY 901 CCGGAGGTTTCCAAATCCTTGAACAGCTGAGAAATCTGGAGAGTGTGGAGTCCCT 960
Db 901 CCGGAGGTTTCCAAATCCTTGAACAGCTGAGAAATCTGGAGAGTGTGGAGTCCCT 960
QY 961 ACAGAGATCTTGGCCGAGAGTCTCAAGCTTAACTCAATTCAGAAATGTTTCCA 1020
Db 961 ACAGAGATCTTGGCCGAGAGTCTCAAGCTTAACTCAATTCAGAAATGTTTCCA 1020
QY 1021 CTGCTACGCTCCGAAAGCTTATGTTCTGGAACAGCTGGGAGAGTTCTGTAAGT 1080
Db 1021 CTGCTACGCTCCGAAAGCTTATGTTCTGGAACAGCTGGGAGAGTTCTGTAAGT 1080
QY 1081 GAAGACCTGGCTCCCAAGATGTTAAAGGCTTTCCCAAGACCGGCTCTCCGCCAGGAA 1140
Db 1081 GAAGACCTGGCTCCCAAGATGTTAAAGGCTTTCCCAAGACCGGCTCTCCGCCAGGAA 1140
QY 1141 GCATTTGTCATGATTAATTTACGCGCCCTGCTCATCTGAGTGTACAGCTTCTGATGAG 1200
Db 1141 GCATTTGTCATGATTAATTTACGCGCCCTGCTCATCTGAGTGTACAGCTTCTGATGAG 1200
QY 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTAAAGCCAGAAATGTGACCTTTTGGCC 1260
Db 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTAAAGCCAGAAATGTGACCTTTTGGCC 1260
QY 1261 TCCCTACAGAAAGGTCCACCAAGCCCAAGTTTACCAATGCTGTGA 1308
Db 1261 TCCCTACAGAAAGGTCCACCAAGCCCAAGTTTACCAATGCTGTGA 1308

RESULT 5
AX772799 1628 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 9 from Patent WO03046167.
DEFINITION AX772799
ACCESSION AX772799
VERSION AX772799.1 GI:32485026
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase
PATENT: WO 03046167-A 9 05-JUN-2003;

/clone="DKFP686K20144"
/issue_type="human colon endothel primary cell culture"
/clone_id="686 (synonym: hlc03). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
3589. .3594
polya_signal
polya_site
3612
ORIGIN

Query Match 95.1%; Score 1243.6; DB 9; Length 3675;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 4; Indels 19; Gaps 2;

QY 1 ATGGGTCAGAGAGCTGTGTGCAAGAGCTGTACAGCTGTGAGAGCTGTCTACATTGTTC 60
DB 51 ATGGGTCAGAGAGCTGTGTGCAAGAGCTGTACAGCTGTGAGAGCTGTCTACATTGTTC 110
QY 61 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 111 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 170
QY 121 AAGCTTACAGAGCTTAAAGAGAGATGTTCCATGATCTTCAATTTCAAGGAGGAGCTT 180
DB 171 AAGCTTACAGAGCTTAAAGAGAGATGTTCCATGATCTTCAATTTCAAGGAGGAGCTT 230
QY 181 CAAGCTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 231 CAAGCTGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290
QY 241 CAGAGAGAGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 291 CAGAGAGAGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 350
QY 301 TCATCTTACTTGAAGCTTGAAGAGGCTGTGAGGAGGCTTATGCGACAGTTTCAAGGGG 360
DB 351 TCATCTTACTTGAAGCTTGAAGAGGCTGTGAGGAGGCTTATGCGACAGTTTCAAGGGG 410
QY 361 ATTAGAGAGATTAATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 411 ATTAGAGAGATTAATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 470
QY 421 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 471 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 530
QY 481 ATTTGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 531 ATTTGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 590
QY 541 ATTCACACAGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 591 ATTCACACAGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 650
QY 601 GTGAGAGCTTTCATGTTTCAACTTTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 651 GTGAGAGCTTTCATGTTTCAACTTTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 710
QY 661 CTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 711 CTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 770
QY 721 GGTGATTTTGT 780
DB 771 GGTGATTTTGT 830
QY 781 GTGAGAGCTCTGT 840
DB 831 GTGAGAGCTCTGT 890
QY 841 CTGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 891 CTGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 950

QY 901 CCTGGGGTTTCCAGATCTTTGAACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 951 CCT-GGGTTTCCAGATCTTTGAACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1009
QY 961 ACAGAGAGATATCTTGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 1010 ACAGAGAGATATCTTGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1069
QY 1021 CTGCTTACGCTTGAAGGCTTTCATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
DB 1070 CTGCTTACGCTTGAAGGCTTTCATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1129
QY 1081 GAAGAGCTGGCCCTCCAGATGCTTAAAGGCTTTCAGAGAGCCGGGCTCCGCCAGGAA 1140
DB 1130 GAAGAGCTGGCCCTCCAGATGCTTAAAGGCTTTCAGAGAGCCGGGCTTCGCCAGGAA 1189
QY 1141 GCAGCTTGTTCATGATATTTATTTTCAAGCCGCTTCATCTGAGCTGTACAGCTTCTGATG 1200
DB 1190 GCAGCTTGTTCATGATATTTTCAAGCCGCTTCATCTGAGCTGTACAGCTTCTGATG 1246
QY 1201 GAGTCTTGTTTTACAGTTTCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
DB 1247 -----GTTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1291
QY 1261 TCCTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308
DB 1292 TCCTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1339

RESULT 7
BC038807
LOCUS
DEFINITION
Homo sapiens amyotrophic lateral sclerosis 2 (juvenile) chromosome
region, candidate 7, mRNA (cdna clone MGC:46456 IMAGE:5200755),
complete cds.
BC038807
BC038807.1 GI:24416556
MGC.
Homo sapiens (human)
SOURCE
ORGANISM
REFERENCES
AUTHORS
Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D.,
Altschuld, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhac, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquelliano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutfard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skaleja, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
JOURNAL MEDLINE
PUBMED
2 (bases 1 to 2305)
REFERENCE
AUTHORS
TITLE
JOURNAL
Strausberg, R.
Direct Submission
Submitted (25-Oct-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaapi, R.,
Madero, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancilpop, S., Thomas, P.J., Touchman, J.W.,
Tauson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

FEATURES
source
1. 2305
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:46456 IMAGE:5200755"
/cissue_type="Brain, adult, 6 pooled whole brains"
/clone_1ib="NIH_MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 2305
/gene="ALS2CR7"
/db_xref="LOCUSID:65061"
198. 1247
/gene="ALS2CR7"
/codon_start=1
/product="ALS2CR7 protein"
/protein_id="AAH38807.1"
/db_xref="GI:24416557"
/db_xref="LOCUSID:65061"
/translation="MTSFHPRGLQAARAKFKSRKPRNSDCFOEEDLRQGFOMKSL
PGAAASYLNLEKLGEGSYATVYKGISRNGQVALKVISMAABEVPPTAIREASL
KGLKXANTVILHDIHTKETLTFFVEYMTDIAOYMSCHPGLHPHNPVLP
LAYITHQVHLHDLKRONLILSHLSELKADRLAKRSIPQOTTSSEVYITMYRPD
ALHGATVSESDIDINGAGCTFIEMFQGFPLFGVSNITELQETKIEWELVGPEDTPG
VSKLPNYPEWPLPTPRSLHVVMNRLGVPBAEDLASOMLKGFPDRVSQAEALVHD
YFSALPSQLYQLPDE"
351. 1205
/gene="ALS2CR7"
/note="S_TKC; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CCD:smart00220"

ORIGIN

Query Match 91.7%; Score 1200; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTCAAGAGCTGTGTCAGAAAGCTGTACAGCCCTGATGAGCTGTGCTACATTTGTTCA 60
DB 45 ATGGGTCAAGAGCTGTGTCAGAAAGCTGTACAGCCCTGATGAGCTGTGCTACATTTGTTCA 104
QY 61 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGAGACACAGGAGGCTGCGTTTC 120
DB 105 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGAGACACAGGAGGCTGCGTTTC 164
QY 121 AAGCTTAAGAGCTTAAGAGCATCATGCTTCCATGACTTCACTTCAACCCAGGAGACTT 180
|||||

DB 165 AAGCTTAAGAGCTTAAGAGCATCATGCTTCCATGACTTCACTTCAACCCAGGAGACTT 224
QY 181 CAAGTCCCGGTCGCCAGAAAGTTCAAGAGTAAGGCCACGAGTAACAGTATTTT 240
DB 225 CAAGTCCCGGTCGCCAGAAAGTTCAAGAGTAAGGCCACGAGTAACAGTATTTT 284
QY 241 CAGGAAGAGATCTGAGGAGGTTTTCAGTGTGAGAGAAAGCTCCCTTTTGGGAGGCC 300
DB 285 CAGGAAGAGATCTGAGGAGGTTTTCAGTGTGAGAGAAAGCTCCCTTTTGGGAGGCC 344
QY 301 TCATCTTACTTGAACCTTGAGAGAGCTGAGGCTCTTATGCAACAGTTTCAAGAGG 360
DB 345 TCATCTTACTTGAACCTTGAGAGAGCTGAGGCTCTTATGCAACAGTTTCAAGAGG 404
QY 361 ATTAGCAGAAATTAATGACCACTAGTGGCTTTTAAAGTCATGAGATGATGAGAGAA 420
DB 405 ATTAGCAGAAATTAATGACCACTAGTGGCTTTTAAAGTCATGAGATGATGAGAGAA 464
QY 421 GGAATCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGAGTTTGAATGCAAT 480
DB 465 GGAATCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGAGTTTGAATGCAAT 524
QY 481 ATTGTCCTCTGATGACATAATCCACAGCAAGAGACCTGACATTTGTTTGAATAC 540
DB 525 ATTGTCCTCTGATGACATAATCCACAGCAAGAGACCTGACATTTGTTTGAATAC 584
QY 541 ATGACACAGACCTGAGCCAGATATATGTCAGCATCCAGAGGCTTATCTCTATAT 600
DB 585 ATGACACAGACCTGAGCCAGATATATGTCAGCATCCAGAGGCTTATCTCTATAT 644
QY 601 GTCAAGCTTTATATGTTTCACTTTTGGGGGCTGCGGTATCATCCACCAACAGCTT 660
DB 645 GTCAAGCTTTATATGTTTCACTTTTGGGGGCTGCGGTATCATCCACCAACAGCTT 704
QY 661 CTTCACAGGAGCTGTAAGACCTGAGAACTCATCATGACCTGAGAGAGCTCAAACTG 720
DB 705 CTTCACAGGAGCTGTAAGACCTGAGAACTCATCATGACCTGAGAGAGCTCAAACTG 764
QY 721 GCTGATTTTGGTCTTGCCCGGAGCAAGTCCATTTCCAGCAGACATCTTTAGAGTC 780
DB 765 GCTGATTTTGGTCTTGCCCGGAGCAAGTCCATTTCCAGCAGACATCTTTAGAGTC 824
QY 781 GTGACCTCTTGATCCGCGCCCTGATGCTTGTGCGGAGCCATGAAATTTCTTGAG 840
DB 825 GTGACCTCTTGATCCGCGCCCTGATGCTTGTGCGGAGCCATGAAATTTCTTGAG 884
QY 841 CTGGAATATGAGGAGGAGGCTGACATCTTTATTTGAAATGTTCCAGGCTCACTTTGTT 900
DB 885 CTGGAATATGAGGAGGAGGCTGACATCTTTATTTGAAATGTTCCAGGCTCACTTTGTT 944
QY 901 CTTGGGGTTTCCAGATCTTTGAAAGCTGAGAAATCTGAGAGGCTGAGAGTCCCT 960
DB 945 CTTGGGGTTTCCAGATCTTTGAAAGCTGAGAAATCTGAGAGGCTGAGAGTCCCT 1004
QY 961 AAGAGATATCTTGCCCGGAGAGTCTTCAAGCTAATCTAATCAATTCAGAAATGTTCCCA 1020
DB 1005 AAGAGATATCTTGCCCGGAGAGTCTTCAAGCTAATCTAATCAATTCAGAAATGTTCCCA 1064
QY 1021 CTGCTAGGCTTGAAGCTTATGCTGAGAAAGGCTGAGAGGCTTCTGAGGCT 1080
DB 1065 CTGCTAGGCTTGAAGCTTATGCTGAGAAAGGCTGAGAGGCTTCTGAGGCT 1124
QY 1081 GAAAGCTGAGCTTCCAGATGCTTAAAGGCTTTTCCAGAGACCGGCTTCCCGCAGGAA 1140
DB 1125 GAAAGCTGAGCTTCCAGATGCTTAAAGGCTTTTCCAGAGACCGGCTTCCCGCAGGAA 1184
QY 1141 GCACTTTTATGATTTATTTTACGCGCTGCTGCAATCTCAGCTTACAGCTTCTGATGAG 1200
DB 1185 GCACTTTTATGATTTATTTTACGCGCTGCTGCAATCTCAGCTTACAGCTTCTGATGAG 1244
RESULT 8
AX772791

LOCUS AX772791 1534 bp mRNA linear PAT 09-JUL-2003
 DEFINITION Sequence 1 from Patent WO03046167.
 ACCESSION AX772791
 VERSION AX772791.1 GI:32485021
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Koehler, R.H.
 TITLE Regulation of human serine/threonine protein kinase
 JOURNAL Patent: WO 03046167-A 1 05-JUN-2003;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source
 1.1534
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 104.1258
 /codon_start=1
 /product="serine/threonine protein kinase"
 /protein_id="CA506249.1"
 /db_xref="GI:32485022"
 /translation="MTSFHPRGLQAARAKFKSKRPSNSDCFQSEEDLRQFQWRKSL
 PRGAASSYLNLEKLGSGSVATYVYKGISRLINGQLVALKVISMAEBSVPTAARESLR
 KGLKHNVIVLHDIHTKELTFVPEYMHQDLQYMSQHPNVALEFVOLFRLG
 LAYTHQHVLRHLKQONILISLGEIKLADFLARAKSI PSQTYSEVVTLMYRPD
 ALIGATEYSELDIWAGCTIFIMFGQPLPFGVSNIVLISLQLEKIMVLEPVEDPMRG
 VSKLPNYPNPMFPLPTPRSLHVMNRLGRVPEADSLASQMLKFPREDRLSADQALVHD
 YFSALPQOLYLPDEBSLFTVSGVRLKPEWCILLASYQKHHPFSCKM"

ORIGIN
 Query Match 90.6%; Score 1185.4; DB 6; Length 1534;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 122 AGCTAAGAGAGCTTAAAGAGCATCATGTTTCATGACTTCATTTACCCGAGGACTTC 181
 DB 72 ATCTAAGAGAGCTTAAAGAGCATCATGTTTCATGACTTCATTTACCCGAGGACTTC 131
 QY 182 AAGTCCCGTCCGCAAGGTTCAAGAGTAAAGGCCAGAGTAAAGTATGTTTC 241
 DB 132 AAGTCCCGTCCGCAAGGTTCAAGAGTAAAGGCCAGAGTAAAGTATGTTTC 191
 QY 242 AGAAGAGAGATCTGAGGAGGTTTTCAGTGAAGAGAGCTCCCTTTGGGGCAGCCT 301
 DB 192 AGAAGAGAGATCTGAGGAGGTTTTCAGTGAAGAGAGCTCCCTTTGGGGCAGCCT 251
 QY 302 CATCTTACTTGAAGCTTGAAGAAAGCTGGGTGAAGCTCTTAAAGGAGTTCACAGGGGA 361
 DB 252 CATCTTACTTGAAGCTTGAAGAAAGCTGGGTGAAGCTCTTAAAGGAGTTCACAGGGGA 311
 QY 362 TTAGAGAGATTAATGAGCAACTAGTGGCTTTAAAGTATCATGATGATGATGATGATG 421
 DB 312 TTAGAGAGATTAATGAGCAACTAGTGGCTTTAAAGTATCATGATGATGATGATGATG 371
 QY 422 GAGTCCCATTTACAGCTATCCGAGAAAGCTTCTCTGTAAGGTTTGAAGTATGCCATA 481
 DB 372 GAGTCCCATTTACAGCTATCCGAGAAAGCTTCTCTGTAAGGTTTGAAGTATGCCATA 431
 QY 482 TTGTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
 DB 432 TTGTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 491
 QY 542 TGCACACAGACCTGGCCAGTATATGTCTACAGATCCAGAGAGGCTTCACTCTCAATAG 601
 DB 492 TGCACACAGACCTGGCCAGTATATGTCTACAGATCCAGAGAGGCTTCACTCTCAATAG 551
 QY 602 TCAGACTTTTCACTTTCACTTTTCCGGGGCTGGCGTACATCCACCAAGCAGTTC 661
 DB 552 TCAGACTTTTCACTTTCACTTTTCCGGGGCTGGCGTACATCCACCAAGCAGTTC 611

QY 662 TTCACAGGAGCTGAAGACCTCAGAACTTACTCATGATCACTTGGGAGAGCTCAAACTGG 721
 DB 612 TTCACAGGAGCTGAAGACCTCAGAACTTACTCATGATCACTTGGGAGAGCTCAAACTGG 671
 QY 722 CTGATTTTGGTCTTGGCCGGGGCAAGTCCATTCACAGCAGACATCACTTTCAGAACTGG 781
 DB 672 CTGATTTTGGTCTTGGCCGGGGCAAGTCCATTCACAGCAGACATCACTTTCAGAACTGG 731
 QY 782 TGACCTCTGTGATCCGCCCCCTGATGCTTTTGTGGAGCCACTGAAATTTCTCTCGAGC 841
 DB 732 TGACCTCTGTGATCCGCCCCCTGATGCTTTTGTGGAGCCACTGAAATTTCTCTCGAGC 791
 QY 842 TGCACATATGAGGTGAGGCTGATCTTTATGAAATGTTCCAGGTCACACTTTGTTTC 901
 DB 792 TGCACATATGAGGTGAGGCTGATCTTTATGAAATGTTCCAGGTCACACTTTGTTTC 851
 QY 902 CTGGGGTTTCCAAATCTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCTTA 961
 DB 852 CTGGGGTTTCCAAATCTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCTTA 911
 QY 962 CAGAGATACTTGGCCGGGAGTCTCCAAAGTACCTAATCAATCCAGAAATGATGCCAC 1021
 DB 912 CAGAGATACTTGGCCGGGAGTCTCCAAAGTACCTAATCAATCCAGAAATGATGCCAC 971
 QY 1022 TGCCCTAGCCCTCGAAGCCTTCATGTTGTCTGAAACAGGCTGGCAGGTTCTGAAAGCTG 1081
 DB 972 TGCCCTAGCCCTCGAAGCCTTCATGTTGTCTGAAACAGGCTGGCAGGTTCTGAAAGCTG 1031
 QY 1082 AAGACTTGGCTCCCAAGTGTAAAGGCTTTCCCAAGACCCGCTCCGCCAGGAG 1141
 DB 1032 AAGACTTGGCTCCCAAGTGTAAAGGCTTTCCCAAGACCCGCTCCGCCAGGAG 1091
 QY 1142 CACTTGTTCATGATTTATTTGAGCCGCTGACATCTCAGCTGTACAGCTTCTGATGAGG 1201
 DB 1092 CACTTGTTCATGATTTATTTGAGCCGCTGACATCTCAGCTGTACAGCTTCTGATGAGG 1151
 QY 1202 AGCTTTTGTTAACAGTTTCAGAGTGAAGGCTTAAAGCCAGAAATGTGACCTTTTGGCT 1261
 DB 1152 AGCTTTTGTTAACAGTTTCAGAGTGAAGGCTTAAAGCCAGAAATGTGACCTTTTGGCT 1211
 QY 1262 CCTACCAAGAAAGTACACCCAGGCCAGTTTAGCAATGCTGTGA 1308
 DB 1212 CCTACCAAGAAAGTACACCCAGGCCAGTTTAGCAATGCTGTGA 1258

RESULT 9
 LOCUS AX772795 1534 bp DNA linear PAT 09-JUL-2003
 DEFINITION Sequence 5 from Patent WO03046167.
 ACCESSION AX772795
 VERSION AX772795.1 GI:32485024
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Koehler, R.H.
 TITLE Regulation of human serine/threonine protein kinase
 JOURNAL Patent: WO 03046167-A 5 05-JUN-2003;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source
 1.1534
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="serine/threonine protein kinase"

ORIGIN
 Query Match 90.6%; Score 1185.4; DB 6; Length 1534;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

122 AGCTAAGAGAGCTAAAGAGAGATGATCTTCATGATCTTCAATTCAGCCAGGAGACTTC 181
 72 ATCTAACAGAGCTAAAGAGAGATGATCTTCATGATCTTCAATTCAGCCAGGAGACTTC 131
 182 AAGCTGCGCGCTCCAGAGAGTCAAGAGTAAAGGCGAGAGTAAAGTGAATGTTTTC 241
 132 AAGCTGCGCGCTCCAGAGAGTCAAGAGTAAAGGCGAGAGTAAAGTGAATGTTTTC 191
 242 AGAAGAGAGATCTGAGAGAGGCTTTCACTGAGAGAGAGCTCCCTTTTGGGAGAGCT 301
 132 AGAAGAGAGATCTGAGAGAGGCTTTCACTGAGAGAGAGCTCCCTTTTGGGAGAGCT 251
 302 CATCTTACTTGAATCTTGAAGAGCTGGGAGAGCTCTTATGAGAGAGTTTACAGAGGA 361
 252 CATCTTACTTGAATCTTGAAGAGCTGGGAGAGCTCTTATGAGAGAGTTTACAGAGGA 311
 362 TTAGAGAGATTAATGAGAGAGCTAGTGGCTTTAAAGTCAATGAGATGAGAGAGAG 421
 312 TTAGAGAGATTAATGAGAGAGCTAGTGGCTTTAAAGTCAATGAGATGAGAGAGAG 371
 422 GAGTCCCATTTACAGATATCCGAGAGCTTCTCTCTGAGAGGCTTGAAGATGCCAATA 481
 372 GAGTCCCATTTACAGATATCCGAGAGCTTCTCTCTGAGAGGCTTGAAGATGCCAATA 431
 482 TTGTGCTCTGATGATGATATCCACAGAGAGAGCTGACATTCGTTTGAATGATA 541
 432 TTGTGCTCTGATGATGATATCCACAGAGAGAGCTGACATTCGTTTGAATGATA 491
 542 TGCACAGAGAGCTGGGCGGAGTATATGCTCAGATCCAGAGAGGCTTCAATCCATATG 601
 492 TGCACAGAGAGCTGGGCGGAGTATATGCTCAGATCCAGAGAGGCTTCAATCCATATG 551
 602 TCAGACTTTTCAATGTTTCACTTTTGGGAGGCTGGGCGTATCCACCAACAGAGCTTC 661
 552 TCAGACTTTTCAATGTTTCACTTTTGGGAGGCTGGGCGTATCCACCAACAGAGCTTC 611
 662 TTCAAGAGAGAGCTGAAAGCTCAAGAGTATCTCATGATCACTGGGAGAGCTCAAGCTGG 721
 612 TTCAAGAGAGAGCTGAAAGCTCAAGAGTATCTCATGATCACTGGGAGAGCTCAAGCTGG 671
 722 CTGATTTTGGTCTTGGCGGCGGAGAGTCAATCCAGAGAGATATCTCTTGAAGAGTGG 781
 672 CTGATTTTGGTCTTGGCGGCGGAGAGTCAATCCAGAGAGATATCTCTTGAAGAGTGG 731
 782 TGACCTCTGATACCGGCGGCGGAGTCTTGGCTGGGAGAGCTGATATCTCTGAGAG 841
 732 TGACCTCTGATACCGGCGGCGGAGTCTTGGCTGGGAGAGCTGATATCTCTGAGAG 791
 842 TGGACATATGGGAGTGGAGCTGATCTTATGAAATGTTTCAAGGATCACTTGTGTTTC 901
 792 TGGACATATGGGAGTGGAGCTGATCTTATGAAATGTTTCAAGGATCACTTGTGTTTC 851
 902 CTGGGGGTTTCAACATCTTGAAGAGCTGAGAGAGAGTGGGAGGCTGGGAGGCTCCCTA 961
 852 CTGGGGGTTTCAACATCTTGAAGAGCTGAGAGAGAGTGGGAGGCTGGGAGGCTCCCTA 911
 962 CAGAGGATATCTTGGCGGAGAGTCTCCAGAGTACTTAATCAATCCAGAGAGTCTCCAC 1021
 912 CAGAGGATATCTTGGCGGAGAGTCTCCAGAGTACTTAATCAATCCAGAGAGTCTCCAC 971
 1022 TGCCTAACGCTTGAAGAGCTTCAATGTTTCTGGAAGAGGCTGGGAGGCTTCTGAAGCTG 1081
 972 TGCCTAACGCTTGAAGAGCTTCAATGTTTCTGGAAGAGGCTGGGAGGCTTCTGAAGCTG 1031
 1082 AAGAGCTGGGCTCCAGAGAGTCTTGAAGAGGCTTCCAGAGAGCGGCTCCGCGCCAGAGAG 1141
 1032 AAGAGCTGGGCTCCAGAGAGTCTTGAAGAGGCTTCCAGAGAGCGGCTCCGCGCCAGAGAG 1091
 1142 CACTTGTTCATGATATTTCAAGGCGCTTGCATCTTCAGCTGATACCAAGCTTCTGATGAG 1201
 1092 CACTTGTTCATGATATTTCAAGGCGCTTGCATCTTCAGCTGATACCAAGCTTCTGATGAG 1151
 1202 AGCTTTTGTTCATGATTTCAAGAGTGAAGGCTTGAAGCAAGATGTTGTGACTTTTGGCCT 1261

Db 1152 AGCTTTTGTTCATGATTTCAAGAGTGAAGGCTTAAAGCAAGATGTTGTGACTTTGGCCT 1211
 Qy 1262 CCTACAGAGAGGATGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308
 Db 1212 CCTACAGAGAGGATGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1258
 RESULT 10
 AB053308 1534 bp mRNA linear PRI 05-OCT-2001
 LOCUS Homo sapiens AL52CR7 mRNA, complete cds..
 DEFINITION
 ACCESSION AB053308
 VERSION AB053308.1 GI:15823641
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 1 Hadano, S., Hand, C.K., Ouga, H., Yanagisawa, Y., Ootomo, A.,
 Devon, R.S., Miyamoto, N., Showguchi-Miyata, J., Okada, Y.,
 Singaraja, R., Figlewicz, D.A., Kwiatkowski, T., Hosler, B.A.,
 Saele, T., Skaug, J., Nasir, J., Brown, R.H.Jr., Scherer, S.W.,
 Rouleau, G.A., Hayden, M.R. and Ikeda, J.-E.
 TITLE A gene encoding a putative GTPase regulator is mutated in familial
 amyotrophic lateral sclerosis 2
 JOURNAL Nat. Genet. 29 (2), 166-173 (2001)
 MEDLINE 21470351
 PUBMED 11586298
 REFERENCE
 2 (bases 1 to 1534)
 AUTHORS Hadano, S. and Ikeda, J.
 TITLE Direct Submission
 JOURNAL Submitted (08-JAN-2001) Shinji Hadano, Tokai University, The
 Institute of Medical Sciences; Bohseidai, Isehara, Kanagawa
 259-1193, Japan (E-mail: shinji@naga.med.u-tokai.ac.jp,
 Tel:81-463-91-5095, Fax:81-463-91-4993)
 FEATURES
 source
 1. 1534
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q33"
 /clone="CAL5124"
 /tissue_type="testis"
 /note="RT-PCR
 product: 1-762nt-clone; IMAGE: 730574; 670-1534nt-composite
 sequence: 1-1534nt"
 1. 1534
 /gene="AL52CR7"
 /gene="AL52CR7"
 104. 1258
 /gene="AL52CR7"
 /codon_start=1
 /protein_id="BAB69017.1"
 /db_xref="GI:15823642"
 /translation="MTSFHPRLQAAQAKRSKRPSNSDFQEBDLRQGFQWRKSL
 PFGAASYLEKLESGSVATVYKGISRLNGQLVALKYSMAAEVGFPTAIRASLL
 KGLKHANVILHDIHTKETLTVEEVMHTDIAOYMSQHPGLHFNHNLFMFOILRG
 LAYITHQVLRDLKRPQNLISLIGELTLARQARARSIPQTSVSVTLVWVPPD
 ALGATREYSEDLTWAGACITFEMGQCPPLFPGVSNITLQLEKIVELVGPREDTWPB
 VSKLPNVNPEWPLPTPSLHVNVNRLGRVPEADLSQMLGPFPRDRVSAQELVHD
 YFSALPSQLYOLPDEBSLFTVSGVRLKPEMDCLASLYQGHHPAFQSKCW"
 ORIGIN
 Query Match 90.6%; Score 1185.4; DB 9; Length 1534;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 122 AGCTAAGAGAGCTAAAGAGAGATGATCTTCATGATCTTCAATTCAGCCAGGAGACTTC 181
 Db 72 ATCTAACAGAGCTAAAGAGAGATGATCTTCATGATCTTCAATTCAGCCAGGAGACTTC 131

QY 182 AAGCTGCCCCGTCGAGAGTTCAAGATGTAAGGCCACGAGATACAGTATGTTTTC 241
DB 132 AAGCTGCCCCGTCGAGAGTTCAAGATGTAAGGCCACGAGATACAGTATGTTTTC 191
QY 242 AAGAAAGGATCTGAGAGGGGTTTTCACTGAGAGAGGCCCTCCCTTTTGGGGCAGCC 301
DB 192 AAGAAAGGATCTGAGAGGGGTTTTCACTGAGAGAGGCCCTCCCTTTTGGGGCAGCC 251
QY 302 CATCTTACTTGAACCTTGAGAGAGCTGGGTGAAGGCTCTTAATGCAAGATTTCAGAGGGA 361
DB 252 CATCTTACTTGAACCTTGAGAGAGCTGGGTGAAGGCTCTTAATGCAAGATTTCAGAGGGA 311
QY 362 TTAGCGAATTAATGAGCACTAGTGGCTTTAAAGTCATCAGCATGAATGCGAGAGAG 421
DB 312 TTAGCGAATTAATGAGCACTAGTGGCTTTAAAGTCATCAGCATGAATGCGAGAGAG 371
QY 422 GAGTCCCATTTACAGCTATCCGAGAGCTCTCTCCGGAAGGGTTGAAACATGCCAATA 481
DB 372 GAGTCCCATTTACAGCTATCCGAGAGCTCTCTCCGGAAGGGTTGAAACATGCCAATA 431
QY 482 TTGTGCTCTGCATGACATATCACAACCAAGAGACCTGACATTGTTTTGAATAGA 541
DB 432 TTGTGCTCTGCATGACATATCACAACCAAGAGACCTGACATTGTTTTGAATAGA 491
QY 542 TGCACACAGACCTGGCCAGATATATGCTCAGACATCCAGAGAGGCTTCACTCATTAAG 601
DB 492 TGCACACAGACCTGGCCAGATATATGCTCAGACATCCAGAGAGGCTTCACTCATTAAG 551
QY 602 TCAGACTTTTCATGTTTCAACTTTTGCGGGGCTGGGGTATCATCCACACACAGCTTC 661
DB 552 TCAGACTTTTCATGTTTCAACTTTTGCGGGGCTGGGGTATCATCCACACACAGCTTC 611
QY 662 TTCACAGGAGCCTGAAACCTCAGAACTTATCTCATGCTCACTGGAGAGCTCAAACTGG 721
DB 612 TTCACAGGAGCCTGAAACCTCAGAACTTATCTCATGCTCACTGGAGAGCTCAAACTGG 671
QY 722 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTTCCACGCCAGACATATCTTCAAGAGTG 781
DB 672 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTTCCACGCCAGACATATCTTCAAGAGTG 731
QY 782 TGACCCCTGCTGATCGGAGCCCTGATGCTTTGCTGGAGGCGCATGATATTTCTCTGAGG 841
DB 732 TGACCCCTGCTGATCGGAGCCCTGATGCTTTGCTGGAGGCGCATGATATTTCTCTGAGG 791
QY 842 TGGACATATGAGGAGTGAAGGCTGATCTTTATGAAATGTTTCAAGGCTCAACTTTGTTTC 901
DB 792 TGGACATATGAGGAGTGAAGGCTGATCTTTATGAAATGTTTCAAGGCTCAACTTTGTTTC 851
QY 902 CTGGGGTTTTCACATCTTTGAAACAGCTGAGAGAAATCTGGAGAGTGTGGAGAGTCCCTA 961
DB 852 CTGGGGTTTTCACATCTTTGAAACAGCTGAGAGAAATCTGGAGAGTGTGGAGAGTCCCTA 911
QY 962 CAGAGGATATTTGGCCGGGAGTCTCCAAAGCTAACCTAACATCCAGAAATGTTTCCAC 1021
DB 912 CAGAGGATATTTGGCCGGGAGTCTCCAAAGCTAACCTAACATCCAGAAATGTTTCCAC 971
QY 1022 TGGCTTACGCTCTGAGAGCTTATGTTTCTGAAACAGGCTGGGCAAGGTTTCTGAAAGCTG 1081
DB 972 TGGCTTACGCTCTGAGAGCTTATGTTTCTGAAACAGGCTGGGCAAGGTTTCTGAAAGCTG 1031
QY 1082 AAGACCTGGGCTCCAGATGCTTAAAGGCTTTTCCAGAGACCGCGTCTCGCCACAGAGAG 1141
DB 1032 AAGACCTGGGCTCCAGATGCTTAAAGGCTTTTCCAGAGACCGCGTCTCGCCACAGAGAG 1091
QY 1142 CACTTGTTCATGATATTTTCAAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1201
DB 1092 CACTTGTTCATGATATTTTCAAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1151
QY 1202 AGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGGCCT 1261
DB 1152 AGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGGCCT 1211
QY 1262 CCAACCAAGAGGTCACCAACCCAGCCAGTTTGAAGCAATGTGTGTGA 1308

DB 1212 CCTACCAAGAAAGGTACCAACCAAGCCAGTTTGAAGCAATGTGTGTGA 1258
RESULT 11
AR265350
LOCUS AR265350 2203 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6492154.
ACCESSION AR265350
VERSION AR265350.1 GI:29693853
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2203)
AUTHORS Yan,C., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6492154-A 1 10-DEC-2002;
FEATURES
Location/Qualifiers
source 1..2203
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 77.1%; Score 1009; DB 6; Length 2203;
Best Local Similarity 100.0%; Pred.No. 4.6e-313; Mismatches 0; Gaps 0;
Matches 1009; Conservative 0; Indels 0; Gaps 0;
QY 1 AAGGGTCAAGAGCTGNGTCAGAAAGCTGACAGCCTGAGTGCAGCTGCTACCATTTGTC 60
DB 118 ATGGGTCAAGAGCTGNGTCAGAAAGCTGACAGCCTGAGTGCAGCTGCTACCATTTGTC 177
QY 61 GAGGAGGCGAGGACACACAGCTGTGAGAGAGTCAAGCTGAGACCAAGAGGCTGCGTTC 120
DB 178 GAGGAGGCGAGGACACACAGCTGTGAGAGAGTCAAGCTGAGACCAAGAGGCTGCGTTC 237
QY 121 AAGCTTACAGACCTTAAAGAACATCATGTTTCCATGACTTCACTTTCACCCCAGGGAGCTT 180
DB 238 AAGCTTACAGACCTTAAAGAACATCATGTTTCCATGACTTCACTTTCACCCCAGGGAGCTT 297
QY 181 CAAGGTGCGCGTCCGACAGAAAGTTCAAGAGTAAAGGCCACGAGATACAGTATGTTT 240
DB 298 CAAGGTGCGCGTCCGACAGAAAGTTCAAGAGTAAAGGCCACGAGATACAGTATGTTT 357
QY 241 CAGAAAGAGATCTGAGGAGGTTTCAAGTGAAGAGAGGCTTCTTTTGGGGCAGCC 300
DB 358 CAGAAAGAGATCTGAGGAGGTTTCAAGTGAAGAGAGGCTTCTTTTGGGGCAGCC 417
QY 301 TCATCTTACTTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTAATGCCACAGTTTAAAGGG 360
DB 418 TCATCTTACTTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTAATGCCACAGTTTAAAGGG 477
QY 361 ATTAGCAAAATTAATGAGCACTAGTGGCTTTAAAGTATCATCAGATGAATGCAAGAGAA 420
DB 478 ATTAGCAAAATTAATGAGCACTAGTGGCTTTAAAGTATCATCAGATGAATGCAAGAGAA 537
QY 421 GGAATCCATTTACAGCTATCCGAGAGGCTTCTCTGGAAGGTTTGAACATGCAAT 480
DB 538 GGAATCCATTTACAGCTATCCGAGAGGCTTCTCTGGAAGGTTTGAACATGCAAT 597
QY 481 ATTGTGCTCTGTCAGATCATATCCACACCAAGAGACATGACATTCGTTTTGAATTC 540
DB 598 ATTGTGCTCTGTCAGATCATATCCACACCAAGAGACATGACATTCGTTTTGAATTC 657
QY 541 ATGCACAGACCTGGCCAGATATATGTTCTCAGATCCAGAGGCTTCACTCATTAAT 600
DB 658 ATGCACAGACCTGGCCAGATATATGTTCTCAGATCCAGAGGCTTCACTCATTAAT 717
QY 601 GTTCAGACTTTTCAATGTTTCACTTTTGGGGCTGGCGTATCATCCACCAACAGCTT 660
DB 718 GTTCAGACTTTTCAATGTTTCACTTTTGGGGCTGGCGTATCATCCACCAACAGCTT 777

Db 358 CAGGAGAGGATCTGAGGCAAGGTTTCACTGGAGGAAAGCCCTTTTGGGGCAACC 417

`/db_xref="taxon:9606"`
`/note="Incyte ID No: 7494145CB1"`

ORIGIN

Query Match 67.9%; Score 887.6; DB 6; Length 1376;
 Best Local Similarity 99.6%; Pred. No. 6.3e-274;
 Matches 890; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 122 AGCTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCACTTTCACCCAGGGGACTTC 181
 DB 266 ATCTTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCACTTTCACCCAGGGGACTTC 325
 QY 182 AAGCTGCGCGTCCGAGAAAGTCAAGAGTAAAGCCAGAGTAACAGTATGTTTTC 241
 DB 326 AAGCTGCGCGTCCGAGAAAGTCAAGAGTAAAGCCAGAGTAACAGTATGTTTTC 385
 QY 242 AGGAAGAGATCTGAGGAGGAGGTTTTCAGTGAAGAAAGCTCCCTTTTGGGGAGCCCT 301
 DB 386 AGGAAGAGATCTGAGGAGGAGGTTTTCAGTGAAGAAAGCTCCCTTTTGGGGAGCCCT 445
 QY 302 CATCTTAAGTGAAGCTTGAAGAAAGTGGTGAAGGCTTTATGAGACAGTTTCAAGAGGA 361
 DB 446 CATCTTAAGTGAAGCTTGAAGAAAGTGGTGAAGGCTTTATGAGACAGTTTCAAGAGGA 505
 QY 362 TTGAGCAATTAATGAGCAATCTAGTGGCTTTAAAGTCAATCAGCATGAATGAGAGGAAG 421
 DB 506 TTGAGCAATTAATGAGCAATCTAGTGGCTTTAAAGTCAATCAGCATGAATGAGAGGAAG 565
 QY 422 GAGTCCCATTTACAGCTATCCGAAAGCTTCTCTCTGAAGGGTTGAAACATGCCAATA 481
 DB 566 GAGTCCCATTTACAGCTATCCGAAAGCTTCTCTCTGAAGGGTTGAAACATGCCAATA 625
 QY 482 TTGTGCTCCCGCAGCATATATCCACCAAGAGACAGCATTCGTTTGAATPACA 541
 DB 626 TTGTGCTCCCGCAGCATATATCCACCAAGAGACAGCATTCGTTTGAATPACA 685
 QY 542 TGCAACAGAGCTGGCCAGATATATGCTCAGCATCCAGAGGGCTTCATCATTAATG 601
 DB 686 TGCAACAGAGCTGGCCAGATATATGCTCAGCATCCAGAGGGCTTCATCATTAATG 745
 QY 602 TCAGACTTTTCAGTGTTCATCTTTTGGGGGCTGGGCTATCCACACAGCAACAGGTTT 661
 DB 746 TCAGACTTTTCAGTGTTCATCTTTTGGGGGCTGGGCTATCCACACAGCAACAGGTTT 805
 QY 662 TTCAACAGAGCTGGAAACCTCAGAACTTCAATCAGTCACTGGGAGAGCTCAAACTGG 721
 DB 806 TTCAACAGAGCTGGAAACCTCAGAACTTCAATCAGTCACTGGGAGAGCTCAAACTGG 865
 QY 722 CTGATTTTGGTCTTGGCCCGGAGCAAGTCAATCCAGAGCAAGATCTTTCAGAAAGTGG 781
 DB 866 CTGATTTTGGTCTTGGCCCGGAGCAAGTCAATCCAGAGCAAGATCTTTCAGAAAGTGG 925
 QY 782 TGACCTCTGATACCGGCGGCTGATGCTTGGTGGAGCCAGCATTAATTTCTTGAGC 841
 DB 926 TGACCTCTGATACCGGCGGCTGATGCTTGGTGGAGCCAGCATTAATTTCTTGAGC 985
 QY 842 TGACATATGGGGGAGGAGGCTGATCTTTATGAAATGTTCAAGGGCACTTTGTTTC 901
 DB 986 TGACATATGGGGGAGGAGGCTGATCTTTATGAAATGTTCAAGGGCACTTTGTTTC 1045
 QY 902 CTGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGAGTCTGGAGTCCCTA 961
 DB 1046 CTGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGAGTCTGGAGTCCCTA 1105
 QY 962 CAGAGATTAATGAGGAGGAGTCTCAGAGTAATCTAATCCAGAAATGGT 1015
 DB 1106 CAGAGATTAATGAGGAGGAGTCTCAGAGTAATCTAATCCAGAGAAATGGT 1159

RESULT 14
 AKI31512
 LOCUS AKI31512 1828 bp mRNA linear PRI 07-MAY-2004
 DEFINITION Homo sapiens cDNA FLJ16732 fis, clone BNGH42005017, moderately similar to SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (BC 2.7.1.-).
 ACCESSION AKI31512

VERSION AKI31512.1 GI:4707537
 KEYWORDS oligo capping, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K., Takehashi, Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1828)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (01-MAR-2004) Takao Isogai, Fuj Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kitarazu, Chiba 292-0818, Japan Fax: 81-438-52-1986 (E-mail: genom@sehiru.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-1986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) (supported by Japan construction: Heijix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..1828
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BNGH42005017"
 /cell_line="H4"
 /cell_type="neurogloma"
 /clone_lib="BNGH42"
 /note="cloning vector: pME18SPU3"
 87..1124
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAD18656.1"
 /db_xref="GI:4707538"
 /translation="MGELCAKTYGPGSCYHCSRGRAGHGRSOPETTYBAFKLTD LKEASCSMTSPHPRGLQARAKQKSKPRNSDCTQEDLRQGTQMKSPRGAASS YLNLKLGEGSYATVYKGISRLNQLVALKYISMAEGSVPTALREASLKLKLAN IYLNLDIHTKETLTFVEFMYLMFMDLRGLAYIHQVLIHRDLKPNMLLSHGLK LADFLARAKSIPSTYSEVVTLMYRPDNLGATSEYSSSLDIWAGACIPFEMFOG PLFPGVSNILRQLKIMWEVLGVPEDTTPGVSKLPNYPGSKQHGWHTDIGQRTA AMOPCSRNVSP"

ORIGIN

Query Match 66.7%; Score 873; DB 9; Length 1828;
 Best Local Similarity 93.8%; Pred. No. 3.3e-269;
 Matches 946; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTCGATGACGTGCTACCATTTGTTCA 60
 DB 87 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTCGATGACGTGCTACCATTTGTTCA 146
 QY 61 GAGGAGGAGGAGGACACAGCTGTGGAGAGAGTCAAGCTGAGACCGAGAGGAGCTGCTTC 120
 DB 147 GAGGAGGAGGAGGACACAGCTGTGGAGAGAGTCAAGCTGAGACCGAGAGGAGCTGCTTC 206
 QY 121 AAGCTAACAGACCTTAAGAAAGCATCATGTTTCATGACTTCACTTTCACCCAGGGAGACTT 180
 DB 207 AAGCTAACAGACCTTAAGAAAGCATCATGTTTCATGACTTCACTTTCACCCAGGGAGACTT 266
 QY 181 CAGAGTCCCGTGGCCAGAGGTTCAAGAGTAAAGGCCAGAGTAAACAGTAAATGTTT 240
 DB 267 CAGAGTCCCGTGGCCAGAGGTTCAAGAGTAAAGGCCAGAGTAAACAGTAAATGTTT 326

QY 241 CAGGAAGAGATCTGAGCAGGGTTTCAATGAGAGAAAGCCTCCCTTTTGGGGCAGCC 300
DB 327 CAGGAAGAGATCTGAGCAGGGTTTCAATGAGAGAAAGCCTCCCTTTTGGGGCAGCC 386
QY 301 TCATCTTACCTGAACTGGGAGAGCTGGGAGAGCTTATATGCAAGTTTCAAGGGG 360
DB 387 TCATCTTACCTGAACTGGGAGAGCTGGGAGAGCTTATATGCAAGTTTCAAGGGG 446
QY 361 ATTAGCAGATTAATGAGCACTAGTGGCTTTTAAAGTATCAGCATGATGAGAGAA 420
DB 447 ATTAGCAGATTAATGAGCACTAGTGGCTTTTAAAGTATCAGCATGATGAGAGAA 506
QY 421 GGAGTCCCTTTACAGCTATCCGAGAGAGCTTCTCTCTGAGAGGTTTGAACATGCCAAT 480
DB 507 GGAGTCCCTTTACAGCTATCCGAGAGAGCTTCTCTCTGAGAGGTTTGAACATGCCAAT 566
QY 481 ATTGCTCTCTGATGATGATTAATCCACACCAAGAGACCTGATTCGTTTGAATAC 540
DB 567 ATTGCTCTCTGATGATGATTAATCCACACCAAGAGACCTGATTCGTTTGAATAC 626
QY 541 ATGCACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCACTCATTAAT 600
DB 627 ATG----- 629
QY 601 GTTCAGACTTTTCACTGTTTCACTTTTGGGGGCTGGCGCTGATCCACCAACACGTT 660
DB 630 -----CTTTTCACTGTTTCACTTTTGGGGGCTGGCGCTGATCCACCAACACGTT 683
QY 661 CTTCACAGGAGACCTGAAACCTGAACTTAATCTCATGATCACTGGAGAGCTCAACATG 720
DB 684 CTTCACAGGAGACCTGAAACCTGAACTTAATCTCATGATCACTGGAGAGCTCAACATG 743
QY 721 GCTGATTTTGTCTTCCCGGCGCAAGTCAATCCAGCAGACATATCTTCAGAGATC 780
DB 744 GCTGATTTTGTCTTCCCGGCGCAAGTCAATCCAGCAGACATATCTTCAGAGATC 803
QY 781 GTGACCTCTTGTGATCCGCGGCTGATGCTTTGCTGGAGACCATGAAATTTCTTGAG 840
DB 804 GTGACCTCTTGTGATCCGCGGCTGATGCTTTGCTGGAGACCATGAAATTTCTTGAG 863
QY 841 CTGGAACATATGAGGGTGCAGAGCTGATCTTATTTGAATGTTCCAGGGTCAACCTTGT 900
DB 864 CTGGAACATATGAGGGTGCAGAGCTGATCTTATTTGAATGTTCCAGGGTCAACCTTGT 923
QY 901 CCTGGGTTTCCAACTCTTTGAACAGCTGAGAGAAATCTGGAGGTTGCTGGAGTCCCT 960
DB 924 CCTGGGTTTCCAACTCTTTGAACAGCTGAGAGAAATCTGGAGGTTGCTGGAGTCCCT 983
QY 961 ACAGAGATATCTTGGCGGAGAGTCTCCAAAGTACTTAACTCAATCAATCCAG 1009
DB 984 ACAGAGATATCTTGGCGGAGAGTCTCCAAAGTACTTAACTCAATCAATCCAG 1032

RESULT 15
AX803407 888 bp DNA 1linear PAT 24-NOV-2003
LOCUS AX803407
DEFINITION Sequence 69 from Patent WO03000901.
ACCESSION AX803407
VERSION AX803407.1 GI:38502070
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1. Martinez, R.A. and Sigurdson, G.T.
Nucleic acids encoding protein kinases
Patent: WO 03000901-A 69 03-JAN-2003;
Decode Genetics EHF. (IS)
Location/Qualifiers
1. 888
/organism="Homo sapiens"

CDS
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
...<1..>888
/note="MOOSE03184"
/codon_start=1
/product="CMGC 1-Cyclin-dependent kinases (CDKs)"
/protein_id="CAE55992.1"
/db_xref="GI:38502071"
/translation="ASSYINLEKLEGGSYATVYKISIRINGQVALKVISMAEEGV
FTAIRBASLKGILKIANIVLADIIHTRETLFVEYVWSYLAQYMSQHPGLHPNV
RLFMQLRGLAYIHQHVILHDLRPMILISLELKLADFLRASISIPQTSSE
VVTLMVRPDLALGATREYSELIDINAGCIPTEMOGOLPFGVSNISLEKIEVL
GVPRDTPGVSKLPVNPDCWRNSIFLSHFRLRVRPAEDLASQMLKGFPRDVS
QDALVHDFSLPSQLYDLP"

ORIGIN

Query Match 62.8%; Score 821; DB 6; Length 888;
Best Local Similarity 95.7%; Pred. No. 1.7e-252;
Matches 858; Conservative 0; Mismatches 30; Indels 9; Gaps 1;

QY 298 GCCTCATCTTACTTGAAGCTTGGAGAGCTGGTGAAGCTCTTATGCGACAGTTTCAAG 357
DB 1 GCCTCATCTTACTTGAAGCTTGGAGAGCTGGTGAAGCTCTTATGCGACAGTTTCAAG 60
QY 358 GGGATTGAGAAATTAATGAGCACTAGTGGCTTTAAAGTATCAGCATGATGAGAG 417
DB 61 GGGATTGAGAAATTAATGAGCACTAGTGGCTTTAAAGTATCAGCATGATGAGAG 120
QY 418 GAAGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCTGAGAGGTTTGAACATGCC 477
DB 121 GAAGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCTGAGAGGTTTGAACATGCC 180
QY 478 AATATTTGCTCTGATGATGATTAATCCACCAAGAGACATGATGCTGTTTGA 537
DB 181 AATATTTGCTCTGATGATGATTAATCCACCAAGAGACATGATGCTGTTTGA 240
QY 538 TACATGACACAGACCTGGCCAGATATATGCTCAGCATCCAGAGGCTTATCCTCAT 597
DB 241 TACATGACAGTATGATGATGATTAATGCTCAGCATCCAGAGGCTTATCCTCAT 300
QY 598 AATGTCAGACTTTTATGTTTCACTTTTGGGGGCTGGCTGATCCACCAACACAC 657
DB 301 AATGTCAGACTTTTATGTTTCACTTTTGGGGGCTGGCTGATCCACCAACACAC 360
QY 658 GTTCTTACAGGAGCTGAAACCTGAGAACTTATCATGATCAGCTGGAGAGCTCA 717
DB 361 GTTCTTACAGGAGCTGAAACCTGAGAACTTATCATGATCAGCTGGAGAGCTCA 420
QY 718 CTGGCTGATTTTGTCTTCCCGGCGCAAGTCAATCCAGCAGACATATCTTCAGAA 777
DB 421 CTGGCTGATTTTGTCTTCCCGGCGCAAGTCAATCCAGCAGACATATCTTCAGAA 480
QY 778 GTGTGACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
DB 481 GTGTGACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 838 GAGCTGACATATGAGGGTGCAGAGCTGATCTTATTTGAATGTTCCAGGGTCAACCTT 897
DB 541 GAGCTGACATATGAGGGTGCAGAGCTGATCTTATTTGAATGTTCCAGGGTCAACCTT 600
QY 898 TTTCTGGGGTTTCCAACTCTTTGAACAGCTGAGAGAAATTTGGAGGTTGCTGGAG 957
DB 601 TTTCTGGGGTTTCCAACTCTTTGAACAGCTGAGAGAAATTTGGAGGTTGCTGGAG 660
QY 958 CCTACAGAGATATGAGCGGAGAGCTCCAAAGTAACTTAACTCAATCAATCAATGTT 1017
DB 661 CCTACAGAGATATGAGCGGAGAGCTCCAAAGTAACTTAACTCAATCAATCAATGTT 720
QY 1018 CCACTGCTTACGCTGAGAGCTTATGTTCTGAGAAAGAGCTGGAGAGGTTCTTGA 1077
DB 721 ACAAACCTTATTTCTCTCC-----ACTTTTCCAGGCTGGAGAGGTTCTTGA 771
QY 1078 GCTGAAGACCTGGCTCCCAAGTGTCTTAAAGGCTTCCCAAGAGACCGCTCTCCGCC 1137

Tue Dec 28 15:07:58 2004

us-10-786-065-4.rge

Page 14

[illegible]

Search completed: December 27, 2004, 04:18:20
Job time : 3575.15 secs

FT	variation	/number= 3 replace(3673,G) /*tag= k
FT	variation	replace(3747,G) /*tag= l
FT	variation	replace(3788,G) /*tag= m
FT	exon	7939..8018 /*tag= n
FT	intron	/number= 4 8019..18260 /*tag= o
FT	variation	/number= 4 replace(8034,T) /*tag= p
FT	exon	18261..18355 /*tag= q
FT	intron	/number= 5 18356..19116 /*tag= r
FT	exon	/number= 5 19117..19179 /*tag= s
FT	intron	/number= 6 19180..29256 /*tag= t
FT	variation	/number= 6 replace(27740,C) /*tag= u
FT	variation	replace(27752,T) /*tag= v
FT	exon	29297..29420 /*tag= w
FT	intron	/number= 7 29421..31091 /*tag= x
FT	variation	/number= 7 replace(29927,T) /*tag= y
FT	variation	replace(30772,G) /*tag= z
FT	exon	31092..31212 /*tag= aa
FT	intron	/number= 8 31213..42825 /*tag= ab
FT	variation	/number= 8 replace(36310,G) /*tag= ac
FT	variation	replace(36337,T) /*tag= ad
FT	variation	replace(40618,C) /*tag= ae
FT	exon	replace(40928,T) /*tag= af
FT	exon	42826..42919 /*tag= ag
FT	intron	/number= 9 42920..50322 /*tag= ah
FT	variation	/number= 9 replace(44701,A) /*tag= aj
FT	variation	replace(46020,C) /*tag= ak
FT	variation	replace(46036,G) /*tag= al
FT	variation	replace(46095,A) /*tag= am
FT	variation	replace(47608,G) /*tag= an
FT	exon	50323..50409 /*tag= ao

FT	/number= 10
FT	replacel(52150,A)
FT	/tag= ap
XX	
PN	
XX	
MO200261060-A2.	
XX	
PD	08-AUG-2002.
XX	
PF	17-JAN-2002; 2002WO-US001106.
XX	
PR	31-JAN-2001; 2001US-026515P.
XX	
PA	09-MAR-2001; 2001US-00801661.
XX	
(PEKE) PB CORP NY.	
PI	
Pt	Yan C, Ketchum K, Di Francesco V, Beasley EM;
DR	
WPI; 2002-608515/65	
P-P9DB; AAO18613, AAO18614.	
XX	
New human kinase peptide and nucleic acid molecule, useful for treating	
PT disorders associated with abnormal expression of kinase protein, e.g.	
PT adenocarcinoma of uterus or lung, in drug screening assays and	
pharmacogenomic analysis.	
PS	
Claim 4; Fig 3; 131pp; English.	
XX	
The present invention provides the protein, cDNA and gene sequences of	
CC two splice variants of a human pteaire family kinase. The sequences are	
CC specifically expressed in the human testis, brain, uterus endometrium	
CC adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and	
CC can be used to treat related diseases. The present sequence is the gene	
of the invention	
CC	
SQ	Sequence 53332 BP; 15976 A; 10335 C; 10952 G; 16069 T; 0 U; 0 Other;
Query Match	100.0%; Score 5000; DB 6; Length 53332;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 5000; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dd	1 TATAGGCGCAATGCCTGGGCTTCAGGCGGTATTCACAGCATTTTGGGAGGAGATCG 60
Oy	61 CTTAGAGCTCAGAATTGGAGACAAGCCTACGTAAATATGTAACCTCTGCTGTACAAA 120
Dd	61 CTTAGAGCTCAGAATTGGAGACAAGCCTACGTAAATATGTAACCTCTGCTGTACAAA 120
Oy	121 TAATPAAAAAATTTTCCAGGCGATGATGCGTGACCCCCAGTGCAGTATTTGGAGGC 180
Dd	121 TAATPAAAAAATTTTCCAGGCGATGATGCGTGACCCCCAGTGCAGTATTTGGAGGC 180
Oy	181 TGAGGTGAGGAATGCTTGAAAGCAGAGGTTGAAGCAAAGCCTAAGGCAACATAGTAGA 240
Dd	181 TGAGGTGAGGAATGCTTGAAAGCAGAGGTTGAAGCAAAGCCTAAGGCAACATAGTAGA 240
Oy	241 CCCGTGTCTATAAAAAATAATTTAGTGTGTCTTGGACAGGCTGACAGTACTACT 300
Dd	241 CCCGTGTCTATAAAAAATAATTTAGTGTGTCTTGGACAGGCTGACAGTACTACT 300
Oy	301 CGAAGACTGAGTGGAGATCACTGAGCCAGAGGCTGAGGCTCAGTGAACAGTGA 360
Dd	301 CGAAGACTGAGTGGAGATCACTGAGCCAGAGGCTGAGGCTCAGTGAACAGTGA 360
Oy	361 TCACCCAGCTGGATTCCAGCCTGGAAAGACAGAGGGAACCTCTGTTTCCAAAAAAAA 420
Dd	361 TCACCCAGCTGGATTCCAGCCTGGAAAGACAGAGGGAACCTCTGTTTCCAAAAAAAA 420
Oy	421 AAAAAAAAAATGCAAGAAAGACATCAATAAATTGAACTGGGAGCACTTAATTATGTATG 480
Dd	421 AAAAAAAAAATGCAAGAAAGACATCAATAAATTGAACTGGGAGCACTTAATTATGTATG 480
Oy	481 AAATTTCACAACTTTTGAAGAAATTAAGACTTTCTGATTAATAAGTATTAATTAATTT 540

Db 481 AATTACAACTTTTAAAGAAATTAAGACTTTCGATTAATAATTAATTAAT 540
Qy 541 ATTATTAATTAATGAATTAATTAATTTCTGAGAACTAGCTTCTGACTCTCAGTTGT 600
Db 541 ATTATTAATTAATGAATTAATTAATTTCTGAGAACTAGCTTCTGACTCTCAGTTGT 600
Qy 601 CAGTCAAACTTTAATAGTCTTTGGCCGGGTGCGGTGCTCAGCGCTGTAATCCGACAC 660
Db 601 CAGTCAAACTTTAATAGTCTTTGGCCGGGTGCGGTGCTCAGCGCTGTAATCCGACAC 660
Qy 661 TTTGGAGGCGGAGGCGGTGATCAAGAGTTAGAGATCGAACCATCTGGCTAA 720
Db 661 TTTGGAGGCGGAGGCGGTGATCAAGAGTTAGAGATCGAACCATCTGGCTAA 720
Qy 721 CGGTGAACCTCTCTCTAATAAAATACAAAAATTAAGCCGGGTGCGGTGCGGAG 780
Db 721 CGGTGAACCTCTCTCTAATAAAATACAAAAATTAAGCCGGGTGCGGTGCGGAG 780
Qy 781 TGTAGTCCAGCTGCTCAGAGAGCTGAGGAGAGATGGTGTGAACCCGGAGGCGAG 840
Db 781 TGTAGTCCAGCTGCTCAGAGAGCTGAGGAGAGATGGTGTGAACCCGGAGGCGAG 840
Qy 841 CTTCGAGTGGCGGAGATGGCGCACTGCACTCAGCGCTGGGAGAGTGGGAGACTG 900
Db 841 CTTCGAGTGGCGGAGATGGCGCACTGCACTCAGCGCTGGGAGAGTGGGAGACTG 900
Qy 901 TCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Db 901 TCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Qy 961 TTTCTTTCTTTTCTTTTCTTTTCAAAAAATTAATTAATTAATTAATTAAT 1020
Db 961 TTTCTTTCTTTTCTTTTCTTTTCAAAAAATTAATTAATTAATTAATTAAT 1020
Qy 1021 TAAGTCTCTCTCTGTCGCAATTTTGAATGAATTAATTAATTAATTAATTAAT 1080
Db 1021 TAAGTCTCTCTCTGTCGCAATTTTGAATGAATTAATTAATTAATTAATTAAT 1080
Qy 1081 CTGAACATTTTAAAAAATCATCATGAGTTCATTAATTAATTAATTAATTAAT 1140
Db 1081 CTGAACATTTTAAAAAATCATCATGAGTTCATTAATTAATTAATTAATTAAT 1140
Qy 1141 GATTAATGACATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1200
Db 1141 GATTAATGACATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1200
Qy 1201 TGTCTCACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Db 1201 TGTCTCACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Qy 1261 AGTGTGCAATTTGAAGACTAGACAGTCAAAACCAATTAAGCAGATTCCTTGACATGT 1320
Db 1261 AGTGTGCAATTTGAAGACTAGACAGTCAAAACCAATTAAGCAGATTCCTTGACATGT 1320
Qy 1321 GTGCTGTAAATAGTACTTTACATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
Db 1321 GTGCTGTAAATAGTACTTTACATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
Qy 1381 CAAGCAAGAAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
Db 1381 CAAGCAAGAAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
Qy 1441 AATATCATTTTGTACAGTTAAGAACATAGTGTCTCAGCAGATTTTGAATTAAT 1500
Db 1441 AATATCATTTTGTACAGTTAAGAACATAGTGTCTCAGCAGATTTTGAATTAAT 1500
Qy 1501 TGAATGAATGGCAACCTTTCTAAGACTAATTTGATTAATTTGAAAGGCAAGAGAT 1560
Db 1501 TGAATGAATGGCAACCTTTCTAAGACTAATTTGATTAATTTGAAAGGCAAGAGAT 1560
Qy 1561 GGAGTGAATTTTGAACCTTTTCTGTTTATTAATTTGAGTCTTGAATTAAT 1620
Db 1561 GGAGTGAATTTTGAACCTTTTCTGTTTATTAATTTGAGTCTTGAATTAAT 1620

Db 1561 GGAGTGAATTTTGAACCTTTTCTGTTTATTAATTTGAGTCTTGAATTAAT 1620
Qy 1621 AAGTATGAGGTAGCAATTTTGTAAAGTCTGAAAAACAATGCTTTTCAATTAAT 1680
Db 1621 AAGTATGAGGTAGCAATTTTGTAAAGTCTGAAAAACAATGCTTTTCAATTAAT 1680
Qy 1681 GGTAGGAGAAAAATTTCTCTCAACATGTCACCTTTAGACAGAGAAACCTTAATTAAT 1740
Db 1681 GGTAGGAGAAAAATTTCTCTCAACATGTCACCTTTAGACAGAGAAACCTTAATTAAT 1740
Qy 1741 TATCACATCAATGATATCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db 1741 TATCACATCAATGATATCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Qy 1801 AAACATATGATACCAATAGCACTAACTTGTGGCCAGAACAGAACCTTAATTTGACAA 1860
Db 1801 AAACATATGATACCAATAGCACTAACTTGTGGCCAGAACAGAACCTTAATTTGACAA 1860
Qy 1861 TTTTATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Db 1861 TTTTATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Qy 1921 TCCCTGTCTGATGTGAGTTTCTTGCACTGATTAAGAAAACTGCTGAAGTTGAGGCT 1980
Db 1921 TCCCTGTCTGATGTGAGTTTCTTGCACTGATTAAGAAAACTGCTGAAGTTGAGGCT 1980
Qy 1981 GCTCCAGGAGAGCCATCATGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 1981 GCTCCAGGAGAGCCATCATGTGATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Qy 2041 AAAGGAGAGAAACAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 AAAGGAGAGAAACAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Qy 2101 CTGGAACCTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
Db 2101 CTGGAACCTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
Qy 2161 GCTGCTACCATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2161 GCTGCTACCATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Qy 2221 CCAAGAGAGCTGCTTCAAGATTTTGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 2221 CCAAGAGAGCTGCTTCAAGATTTTGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Qy 2281 AACCAAGAGATTCAGACCTCTCTTTTGTAGCGGAGATCTTCTGCGGTAGGCT 2340
Db 2281 AACCAAGAGATTCAGACCTCTCTCTTTTGTAGCGGAGATCTTCTGCGGTAGGCT 2340
Qy 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2400
Db 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2400
Qy 2401 TTTGGGGAATTTCTTGTAAACCAAAAGGGAATAATCTTGGCTTGGCTGACGAA 2460
Db 2401 TTTGGGGAATTTCTTGTAAACCAAAAGGGAATAATCTTGGCTTGGCTGACGAA 2460
Qy 2461 ACTCACTGGCTTGAAGTGAAGAAAGTGTCTCAAAATCTCAAGGTCTTAATTAAT 2520
Db 2461 ACTCACTGGCTTGAAGTGAAGAAAGTGTCTCAAAATCTCAAGGTCTTAATTAAT 2520
Qy 2521 AAGAGTGAACCTTAAGGCAAGCTGACATTAATTAATTAATTAATTAATTAATTAAT 2580
Db 2521 AAGAGTGAACCTTAAGGCAAGCTGACATTAATTAATTAATTAATTAATTAATTAAT 2580
Qy 2581 TATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
Db 2581 TATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
Qy 2641 AGTGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
Db 2641 AGTGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700

QY 2701 AGAGATTAGTTAACTGATGTTAAAGACACATTTTGTGTGAAGAAAGATGTAGTA 2760
DB AGAGATTAGTTAACTGATGTTAAAGACATTTTGTGTGAAGAAAGATGTAGTA 2760
QY 2761 TTTTCAAGAGTGTATACATTTATGTGCTTTTTCAGGGATCTTTCAAGAAAAGTCCCTTTG 2820
DB TTTTCAAGAGTGTATACATTTATGTGCTTTTTCAGGGATCTTTCAAGAAAAGTCCCTTTG 2820
QY 2821 GGGGTACAGGAAGCTTAGAAAACATTTGAAAGATGAAAATGAGCAAAATTAAGAAAAAT 2880
DB GGGGTACAGGAAGCTTAGAAAACATTTGAAAGATGAAAATGAGCAAAATTAAGAAAAAT 2880
QY 2881 GGTTTTACCAAGGACATGGAATCTTTACTTTGCAATTAATTTTATTTGCTCTTTCTTTT 2940
DB GGTTTTACCAAGGACATGGAATCTTTACTTTGCAATTAATTTTATTTGCTCTTTCTTTT 2940
QY 2941 CTCTAGCTAACAGACCTTAAGAAAGCATGTTCCATGACTTCAATTTCCACCCAGGGGA 3000
DB CTCTAGCTAACAGACCTTAAGAAAGCATGTTCCATGACTTCAATTTCCACCCAGGGGA 3000
QY 3001 CTTCAGCTGCCCCGTGCCGAAAGTTCAAGAGTAAAGGCGACGAGTAACTGATTTGT 3060
DB CTTCAGCTGCCCCGTGCCGAGAGTTCAAGAGTAAAGGCGACGAGTAACTGATTTGT 3060
QY 3061 TTTTCAAGAAAGAGATCTGAGGCGAGGTTTCACTGGGTGATGAGACAGCTGATTTGATC 3120
DB TTTTCAAGAAAGAGATCTGAGGCGAGGTTTCACTGGGTGATGAGACAGCTGATTTGATC 3120
QY 3121 AAGAAATTTTAAATGTGAGCTTGTCTACGGAAGGCGGCTCTTCAAGGCAATTAAT 3180
DB AAGAAATTTTAAATGTGAGCTTGTCTACGGAAGGCGGCTCTTCAAGGCAATTAAT 3180
QY 3181 GAGCGAGCTTCCCAAGTCTGCTGCGCAATGCTGTAAATTTCCCTGGGAAAAAAGT 3240
DB GAGCGAGCTTCCCAAGTCTGCTGCGCAATGCTGTAAATTTCCCTGGGAAAAAAGT 3240
QY 3241 CAACACTTAAAAAAGTGTCTTCTCTCTCCCTTTCACCCGCTCTTTCCCAATTC 3300
DB CAACACTTAAAAAAGTGTCTTCTCTCTCCCTTTCACCCGCTCTTTCCCAATTC 3300
QY 3301 CCTAGAGCAGAGAAAGCTCCCTTTTGGGCGAGCTCATCTTACTTGAATCTGAGAA 3360
DB CCTAGAGCAGAGAAAGCTCCCTTTTGGGCGAGCTCATCTTACTTGAATCTGAGAA 3360
QY 3361 GCTGGGTGAAGCTCTTATGCGACAGTTTAAAGGGATTTAGCAGGTGAGTGAACAATAG 3420
DB GCTGGGTGAAGCTCTTATGCGACAGTTTAAAGGGATTTAGCAGGTGAGTGAACAATAG 3420
QY 3421 CTGGGAGAGCTTTAGATGAGATGAGTCCGCCCCCAATTTCAATTAATAAGCAGGT 3480
DB CTGGGAGAGCTTTAGATGAGATGAGTCCGCCCCCAATTTCAATTAATAAGCAGGT 3480
QY 3481 GAGACATCATAGAAGTTCATAGCACTCAGGACCTGTGCAAGACACATGCGCAGAGGA 3540
DB GAGACATCATAGAAGTTCATAGCACTCAGGACCTGTGCAAGACACATGCGCAGAGGA 3540
QY 3541 GAGAGCATGTATTAATTTAAACAGCTTTGAAAAGAAAACAACTGCGCCCTTAATTA 3600
DB GAGAGCATGTATTAATTTAAACAGCTTTGAAAAGAAAACAACTGCGCCCTTAATTA 3600
QY 3601 AATCAGCCCACTTAATGTTTATCAGCTTCCCTTTGCAATTCATTAAGAGAAATTA 3660
DB AATCAGCCCACTTAATGTTTATCAGCTTCCCTTTGCAATTCATTAAGAGAAATTA 3660
QY 3661 AAGAAATAGACATTTCTACTACTGACCCAAAGAACATTAATCACTTTCAGGCGCTGTG 3720
DB AAGAAATAGACATTTCTACTACTGACCCAAAGAACATTAATCACTTTCAGGCGCTGTG 3720
QY 3721 GAGGCGACAGTTGGTAAAGGCTCTTAACAGTTTTTTTATATCCCTCCTTAATCACAAT 3780
DB GAGGCGACAGTTGGTAAAGGCTCTTAACAGTTTTTTTATATCCCTCCTTAATCACAAT 3780
QY 3721 GGAGGCAAGTGTGTAAAGGCTCTTAACAGTTTTTTTATATCCCTCCTTAATCACAAT 3780

QY 3781 GACAGATTTTGTATATGCGAACCTGGAATTTGCTGCTCATTTCTCCACCTGCGCTTAT 3840
DB GACAGATTTTGTATATGCGAACCTGGAATTTGCTGCTCATTTCTCCACCTGCGCTTAT 3840
QY 3841 AAGAAACTGAAGTTGGTTTTCTGCAAAATTATGTATGACATGCAAAAGATGATTAATCCTAG 3900
DB AAGAAACTGAAGTTGGTTTTCTGCAAAATTATGTATGACATGCAAAAGATGATTAATCCTAG 3900
QY 3901 ATTTTATATTTTGCAGAAATACACAAATGCTCGAGAAATTAATCTGCTTATCCAA 3960
DB ATTTTATATTTTGCAGAAATACACAAATGCTCGAGAAATTAATTAATGCTTTATCCAA 3960
QY 3961 AGCTAAGTACTAATTTGGTAAACCAACACTTTGTAAATATATGTATAAGATCAATGA 4020
DB AGCTAAGTACTAATTTGGTAAACCAACACTTTGTAAATATATGTATAAGATCAATGA 4020
QY 4021 ATTCCTCTTTTATGTCAGAGTGGGAAAGTTGAGTGGTGGTTTTTCTTTATGTATCTCCA 4080
DB ATTCCTCTTTTATGTCAGAGTGGGAAAGTTGAGTGGTGGTTTTTCTTTATGTATCTCCA 4080
QY 4081 ATAGAGAGAAAAGTATGCTCAATAGTGTAAATATTAATTTTAAATATAGCTGAT 4140
DB ATAGAGAGAAAAGTATGCTCAATAGTGTAAATATTAATTTTAAATATAGCTGAT 4140
QY 4081 ATAGAGAGAAAAGTATGCTCAATAGTGTAAATATTAATTTTAAATATAGCTGAT 4140
DB ATAGAGAGAAAAGTATGCTCAATAGTGTAAATATTAATTTTAAATATAGCTGAT 4140
QY 4141 CCGAGTGCAGTGGTGTATCAACTACTTGATCACAACAGTTACAGATTTCTTGTGCT 4200
DB CCGAGTGCAGTGGTGTATCAACTACTTGATCACAACAGTTACAGATTTCTTGTGCT 4200
QY 4201 TCTCCACTCCCACTGCTTCACTTAACTGCGCAAAAACGAAAAAGAAAAATTTATATA 4260
DB TCTCCACTCCCACTGCTTCACTTAACTGCGCAAAAACGAAAAAGAAAAATTTATATA 4260
QY 4261 CTACTCAAGACTAAATATTTATTTATCTAGTATTTATAGCTGTATTTATTTT 4320
DB CTACTCAAGACTAAATATTTATTTATCTAGTATTTATAGCTGTATTTATTTT 4320
QY 4261 CTACTCAAGACTAAATATTTATTTATCTAGTATTTATAGCTGTATTTATTTT 4320
DB CTACTCAAGACTAAATATTTATTTATCTAGTATTTATAGCTGTATTTATTTT 4320
QY 4321 ACTTGTAAACAGATTTGAGGAGACATACAGTTTATTTTATTTATTTATTTAT 4380
DB ACTTGTAAACAGATTTGAGGAGACATACAGTTTATTTTATTTATTTATTTAT 4380
QY 4381 TTAATTTATTTTGAATGGAATCTGTGACCCACGCTGAGTGCAGTGTGATC 4440
DB TTAATTTATTTTGAATGGAATCTGTGACCCACGCTGAGTGCAGTGTGATC 4440
QY 4441 TCGATGATCGAACCTCGCTCGAGTTCAACCACTCGCCCTGAGCCCTTTA 4500
DB TCGATGATCGAACCTCGCTCGAGTTCAACCACTCGCCCTGAGCCCTTTA 4500
QY 4441 TCGATGATCGAACCTCGCTCGAGTTCAACCACTCGCCCTGAGCCCTTTA 4500
DB TCGATGATCGAACCTCGCTCGAGTTCAACCACTCGCCCTGAGCCCTTTA 4500
QY 4501 TACTTCTTAATCTGTTTTAGTCATGAGTACCTTAACCTTTTCAATGCTGAGAACATC 4560
DB TACTTCTTAATCTGTTTTAGTCATGAGTACCTTAACCTTTTCAATGCTGAGAACATC 4560
QY 4561 TGCATTAAGGACACATTTTATTTTATCTAAGCTTCTCATATCAATTTGCGCATGT 4620
DB TGCATTAAGGACACATTTTATTTTATCTAAGCTTCTCATATCAATTTGCGCATGT 4620
QY 4621 AACTGTTTTCAAGTGGGCTCGGAAACGGGGGACCCCTGGAACATCTGGAATCAATGGGCA 4680
DB AACTGTTTTCAAGTGGGCTCGGAAACGGGGGACCCCTGGAACATCTGGAATCAATGGGCA 4680
QY 4681 CCATGGAACATCTGATCCTCTCTCTGAGTTCGACATTTGATGTTGTGCAAGACCTT 4740
DB CCATGGAACATCTGATCCTCTCTCTGAGTTCGACATTTGATGTTGTGCAAGACCTT 4740
QY 4741 TCCAGCCGGAAGTTTACAGAAATTCATTAATCTTTTCTTCTGAGTTACTTATAGTTTTCT 4800
DB TCCAGCCGGAAGTTTACAGAAATTCATTAATCTTTTCTTCTGAGTTACTTATAGTTTTCT 4800
QY 4801 TTTTCAATTAATCTTTTCACTGAGGAAATTAATGATTAATCAACATCAACCCAGC 4860
DB TTTTCAATTAATCTTTTCACTGAGGAAATTAATGATTAATCAACATCAACCCAGC 4860
QY 4861 TCCATTTATTAATTTTCTTTTCTGATGTTGAAATGCTGCATGATTAATTAATAG 4920

DB 4861 |||||
TCATTATTAGATTTCTTCTCGAGGCTTGAAGCTCCAGATTATATTAG 4920
QY 4921 ATCTCAGAACTTGAATCTTCTGTTCAATCTTTAAATCAGTTCTTAAT 4980
|||||
DB 4921 ATCTCAGAACTTGAATCTTCTGTTCTTAATCTTTAAATCAGTTCTTAAT 4980
|||||
QY 4981 CTAATCTTTCTTAATTTGT 5000
|||||
DB 4981 CTAATCTTTCTTAATTTGT 5000
|||||
RESULT 2
AAK71358/c
ID AAK71358 standard; DNA; 37314 BP.
XX
XX AAK71358;
XX
XX 06-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26170.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

PR 24-APR-2001; 2001US-0286591P.
XX (GENE-) GENE LOGIC INC.
PA (PROC) PROCTER & GAMBLE CO.
XX JI D, Axelrod DW, Cook JS, Jalawal N, Einstein R, Houghton A;
PI Metz L;
XX WPI; 2002-557663/59.
XX
XX Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
XX Claim 1; SEQ ID NO 114; 78pp + Sequence Listing; English.
XX
XX The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated CDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 172570 BP; 45194 A; 43991 C; 41126 G; 42258 T; 0 U; 1 Other;
SQ

Query Match 7.4%; Score 370.6; DB 6; Length 172570;
Best Local Similarity 68.2%; Pred. No. 1.9e-51;
Matches 639; Conservative 0; Mismatches 249; Indels 49; Gaps 7;

QY 11 TCGTGTGCTCAAGCGGTATTCCTCCAGCACTTTGGAGAGC-----AGAGAGTCGCTTG 64
DB 136255 TCGGTGTGCTCAAGCGGTATTCCTCCAGCACTTTGGAGAGC-----AGAGAGTCGCTTG 136314
QY 65 AGCTCAGAAATTGAGACAGCGCTACGTAAATAGTGAACCTCTGTCTGTACAAATAAT 124
DB 136315 AGGTGAGAGATTGAAACAGCGCTGACCAATGTTGAACG-CTGTCTGTACAAATAAT 136373
QY 125 AAAAAGATTTTCCAGGATGTGTGGCTGACCCCGAGTGTCTATTGGAGAGCTGAG 184
DB 136374 AAAAAAATTAGCCAGGCGGTGTGGCTGACCACTGTATCCAGCACTATGGAGGCGCGAG 136433
QY 185 GTAGAGAGAAATGCTTGAAGCCAGAGTTGAAGCAAGCCTAAGCAATATGAGACCT 244
DB 136434 ATGGGTGATCTCTTGAAGTCAAGAGTTATGACACCGCTGCGCAATATGAGACCT 136493
QY 245 GTGTCTATTAATA-----AATTAATTAAGTGTGTCTTGGACAGCGCTGCA--GCTAGCTA 298
DB 136494 ATCTCTACTAATAATAATAATAATTAAGCGGCGGTGTGGAGCACTGTAAATCCAGCTA 136553
QY 299 CTCGGAAGCTGAGGTGGAGAG--TCACTGAGCCCGAGAGGCTGAGGCTGCAATGAACG 357
DB 136554 CTCGAGAGGCTGAGGAGAGAAATGTTGAACCCCGAGAGGTGAGAGTTCAGGTGAGCA 136613
QY 358 TGATCACTCAGCTGATTT-----CCAGCCTGGAAGACAGAGGAGACCTGTTTCAAAA 412
DB 136614 AGATCACTCAGCTGATTCAGCTGAGCTGAGGCGCGAGAGCAAGCTCATCTCAAAA 136673
QY 413 AAAAAAAAAAAAAAAAAATGCAAGAAAAAGACATCAATTAACCTTGACTGGAGCATTAACCTTTT 472
DB 136674 AAAAAAAAAAACAACAATAAAAAAAAACAATTAGCAGAGCATATGTGTGCTACAGAGACTGA 136733
QY 473 ATGTGATGAATTCACAACTTTTGTAGGAAGAAATTAGCATTTCTGATAAATGATATA 532

DB 136734 GGCAGAGAAATCACTTGAACCCAGAGAGGCGAGGTTGCACTAGCCAAATGTGCACT 136793
QY 533 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 592
DB 136794 GCATTCAGAGCTGGGTGACAGATTAAGTCCGCTCAAAAA----- 136836
QY 593 TCAAGTTTCAGTCAAAATCTTAATGTCTTTGCGCGGTGGGTGCTCAAGCCTGTAT 652
DB 136837 -----AAAAAAAAAAAAAAAAATTTTGGCAGGCGGTGTGCTCAAGCCTGTAT 136883
QY 653 CCAGCACTTTGGAGAGCGCGAGGCTGATCAACAGTTAGAGATGAGACCATCT 712
DB 136884 CCCCAGCTTTGGAGAGCGCAAGCGGCGGATCAAGAGTCAAGAGTGAACCATCT 136943
QY 713 GGTAAACAGGTGAACCTGTCTTACTTAAATAAATAAATAATTAAGCGGCTGCTGC 772
DB 136944 AGCTAAACAGGTGAACCCCGCTCTCTACTTAAATAAATAAATAATTAAGCGGCTGCTGC 137003
QY 773 CAGAGCGCTGTAGTCCAGCTGCTCAAGAGGCTGAGGAGAAATGTGTGAACCCGGG 832
DB 137004 CAGGCGCTGTAGTCCAGCTGCTCAAGAGGCTGAGGAGAAATGTGTGAACCCGGG 137063
QY 833 AGCGGAGCTTGCAGTGAAGCCAGATTGGCGCACTGCACTCCAGCCTGGCGACAGTGC 892
DB 137064 AGCGGAGCTTGCAGTGAAGCCAGATTGGCGCACTGCACTCCAGCCTGGCGACAGTGC 137123
QY 893 AGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTTGA 929
DB 137124 AGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAATTA 137160

RESULT 4
AAK76548/C
ID AAK76548 standard; DNA; 13673 BP.
XX
XX AAK76548;
AC
XX 07-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31360.
DE
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW
XX cytoelastic; gene therapy; vaccine; metacastis; ds.
XX
XX Homo sapiens.
OS
XX
XX MO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001354.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.

Sequence 13673 BP: 3494 A; 2746 C; 3372 G; 4061 T; 0 U; 0 Other;

Query Match 7.0%; Score 351.8; DB 4; Length 13673;
Best Local Similarity 67.1%; Pred. No. 2e-48;
Matches 626; Conservative 0; Mismatches 282; Indels 25; Gaps 8;

13 CTGTGCTCAGCGGTGATTTCCAGACATTTGGAGGCG-----AGGAGATCGCTGAG 66
DB CAGTGGCTCAGCTGTATATCCAGACATTTGGAGGCGCAAGGTGGTGTATCTACTGTAG 7827
QY 67 CTGAGGAATGGAGCAAGCCTACGTATAGTGAACCTCTGTCTGTACAAATATATA 126
DB 7826 GTGAGAGTTAGAGCCAGCCTGGCCAAAGGGTGAACCCATCTCTATATAATATA 7767
QY 127 AAGAAATTTCCAGGATGTGGCGTGCACCCCAAGTGCAGCTATTTGGAGGCTGAGGT 186
DB 7766 AAAATTAGGCCAGGATGTGGCTGCACCTGTATCCAGACATTTGGAGGCGCAGAGC 7707
QY 187 AGGAGGAATGTTGAAGCCAGAGTGAAGACAGCCTAGGCAATATGTAGACCTTGT 246
DB 7706 AGGTG-ATATCAGAGGTGAGAGATCAAGACCATCTGCTTAACAGGTGAACCCCTGT 7649
QY 247 GTCTATA-----AAAAATAATAGCTGTTGTCTTGGCAGAGGCTGAG--CTAGCT 297
DB 7648 TCTACTAAAAATACAAAAAAATTTAGCCGGCATGTGACAGCGCTGTATCCAGCT 7589
QY 298 ACTCGGAAGCTGAGGTGGAGGATCAC-TGAGCCAGAGGCTGAGGCTGAGTGAACA 356
DB 7588 ACTTAGAGGCTTAAGCAGAGGAGATCACTTGAACCTGAGAGGAGAAAGCTGAGGGAGCC 7529
QY 357 GTGATCACCAGCTGATTTCCAGCTGGAAGACAGAGGAGACCTGTTTCCAAAAAAA 416
DB 7528 GAGATCAGCCCATACACTCCAGCCTGGGCGACAGAGGAGACTGTGTACCAAAAAAAA 7469
QY 417 AAAAAAAAAAATGCAAGAAAGACATATATACTTGACCTGGAGCATATCTTTATGT 476
DB 7468 AAAAAAAAAAGAAAGCTGATTTAGATTTGAACCAAGTTCACTTATATTAAGTAC 7409
QY 477 GATGAATTCACAAATCTTTAGAGAAATTTAGCAATTTGTATATAATATATTA 536
DB 7408 AAAACGGTTACTTCAGTATAGTCAAAAGTCAAAATTTCAATTAATAGACAGGCCACACA 7349
QY 537 TATATTAATTAATTAAGAAATTAATTTCTGAGAAATAGCTTCA-----CTCTC 592
DB 7348 AAAAGAAACAAAGGGGATCATTTGAGCAAAATGATGCCAAGCACTTCAACTGTCTAT 7289
QY 593 TCAGTTGTCAGTCAAAATCTTATAGGTCTTTGGCGGGTGGCGGTGCTCAGCCCTGTAT 652
DB 7288 TCAATTTCTCTTAAGAAATTTCTTCTGCTGGTGGTGGCTGACCGCTGTAT 7229
QY 653 CCCAGCACTTTGGAGGCCGAGGCGGGTGAATCACAAAGTTAGAGATGAGACCAATCT 712
DB 7228 CCCAGCACTTTGGAGGCCGAGGCGGGTGAATCATGAATGACAGAGATCAAGCCATCT 7169
QY 713 GGCTTAACAGGTGAACCTCTCTCTACTATAAAATACAAAAAATTTAGCCGGTGGCTGC 772
DB 7168 GGCTTAATAGGTGAACCAATCTCTACTATAAAATAC-AAAAACAAATTTAGCCAGGTG 7110
QY 773 CAGAGCCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGGAGAAATGTGTAA--CCCG 830
DB 7109 TGGGGCCCTGTAGTCCAGCTACTGGAGGCTGAGGAGGAGAAATGTGTAAAGCTTCCG 7050
QY 831 GGAGCGGAGCTTGAAGTGAAGCCAGATTCGCCCACTGCACTCCAGCCTGGAGCAGTGTG 890
DB 7049 GGAGGTGAGCTTGAAGGAGCTGAGATGTGCACTCCAGCTCCAGCCTGGAGCAGTGTG 6990
QY 891 CGAGACTCTGTCAAAAAAAAAAAAAAAAAAAAA 923
DB 6989 CAAGACTTGTCTCAAAAAAAAAAAAAAAAAAAAA 6957

AAK81195 standard; DNA; 13673 BP.

AC AAK81195; 13673 BP.
XX 07-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36007.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
PN WO200157182-A2.
PD
XX 09-AUG-2001.
PF
XX 17-JAN-2001; 2001WO-US001354.
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 07-JUN-2000; 2000US-0205515P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0228267P.
PR 14-AUG-2000; 2000US-0228268P.
PR 14-AUG-2000; 2000US-0228270P.
PR 14-AUG-2000; 2000US-0228271P.
PR 14-AUG-2000; 2000US-0228275P.
PR 14-AUG-2000; 2000US-0228279P.
PR 18-AUG-2000; 2000US-0228279P.
PR 22-AUG-2000; 2000US-0228681P.
PR 22-AUG-2000; 2000US-0228686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227093P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241856P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 36007; 3071bp + Sequence Listing; English.

XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (1)
XX amino acid sequences given in AAM62170 to AAM91921. (1) have cytosolic
XX activity, and can be used in gene therapy and vaccine production. (1)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (1) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (1) by expressing inactive proteins or to
XX supplement the patient's own production of (1). Additionally, (1)
XX polynucleotides may be used to produce the secreted (1), by inserting the
XX nucleic acids into a host cell and culturing the cell to prevent the
XX protein. (1) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM62169
XX represent sequences used in the exemplification of the present invention

XX Sequence 13673 BP; 3494 A; 2746 C; 3372 G; 4061 T; 0 U; 0 Other;

Query Match 7.0%; Score 351.8; DB 4; Length 13673;
Best Local Similarity 67.1%; Pred. No. 2e-48;
Matches 626; Conservative 0; Mismatches 282; Indels 25; Gaps 8;

QY 13 CTGTGCTACCGCGTATATCCAGACCTTGGGAGGC-----AGAGGATCGCTGAG 66
DB 7886 CAGTGGCTCACCTGTATATCCAGCACTTGGGAGGCCAAGGTGGTGATCACTTGAG 7827
QY 67 CTCAGGATTTGAGACAAGCCTACGTAACATATGTAACCTCTGTACAAATATTA 126
DB 7826 GTTCAGAGTTAAGACACGCTCGGCAACAGGTGTAACCCATCTCTAATAAATAGA 7767
QY 127 AAGATTTTTCAGGCAAGTGTGCGTGACCCCACTGCCAGTATTTGGAGGCTGAGGT 186
DB 7766 AAAATTAGGCCAGGAGTGTGCTCACCTTATATCCAGACCTTGGAGGCGAGGC 7707
QY 187 AGAGGAATGCTTGAAGCAGGAGTTGAAGACAAGCCTAGGACAATAGTGAACCTGT 246
DB 7706 AGGTG--AATCAAGAGTGAAGATCAAGACATCTGCTTAACACGTTGAACCCCT 7649
QY 247 GTCCTATA-----AAAAATAATTAGCTGTGTCTTGGACAAGGCTTGCAG--CTAGCT 297
DB 7648 TCTACTAATAAATACAAAAAATTAGCCGGAGATGTTGACAGCGCTGTATGCCAGCT 7589
QY 298 ACTCGAAGACTGAGTGGAGAGATAC--TGAGCCAGAGGCTGAGGCTGACATGAAGA 356
DB 7588 ACTTAGAGGCTTAAGGCAAGAAATCACTTGAACCTGAAGGACAAGCTGAGGAGGC 7529
QY 357 GTGATCACCCAGTGTGATTCAGCCCTGGAAGACAGAGGAGACCCCTGTTCCAAAAA 416
DB 7528 GAGATCACCGCACTACCTCAGCCTGGCGGACAGAGGAGACTCTGTCAACCAAAAAA 7469

```
QY 417 AAAAAAAAAAATGCAAGAAAGACATCAAACTTACCTGGAGCATATCTTTATGT 476
DB 7468 AAAAAAAAAAAGAGCTGTATTAATTTGAAACCACTTCACTCATTATTAAGTAC 7409
QY 477 GATGAATTCACAACTTTTGAAGAAATAGCATTTCTGATTAATTAATTAATTA 536
DB 7408 AAAACGGTACTCAGCTGATGATCAAAAGTCAAACTTCAATTAATTAATTAATTA 7349
QY 537 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 592
DB 7348 AAAAGAAACAAAAGGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 7289
QY 593 TCAGTTCAGTCAAACTTAAATGATCTTGGCCGGGAGGAGGAGGAGGAGGAGGAGG 652
DB 7288 TCATTTCTCTTAAGAAATTTTCTCTCTGTTGGGAGGAGGAGGAGGAGGAGGAGG 7229
QY 653 CCCAGCACTTTGGAGAGCCGAGGCGGAGTGAATCAAGATTAGAGATCGAGACATCT 712
DB 7228 CCCAGCACTTTGGAGAGCCGAGGCGGAGTGAATCAAGATTAGAGATCGAGACATCT 7169
QY 713 GGTAAACAGGTGAAACCTGCTCTACTTAAATTAATTAATTAATTAATTAATTAAT 772
DB 7168 GGTAAATACGCTCAACACCATCTCTACTTAAATTAATTAATTAATTAATTAATTA 7110
QY 773 CAGACGCTGTAGTCCAGCTGCTCAAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGG 830
DB 7109 TGGGCGCTGTAGTCCAGCTGCTCAAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGG 7050
QY 831 GGAGGCGAGCTTGCAGTGAAGCCGAGATTGGCCCACTGCACTCAGCTGGGAGAGT 890
DB 7049 GAGAGGAGGAGCTTGCAGGCGAGCTGAGATGCTGCACTGCACTCAGCTGGGAGAG 6990
QY 891 CGAGACTCTGTCTCAAAAAAAAAAAAAAAAAA 923
DB 6989 CAGACTTGTCTCAAAAAAAAAAAAAAAAAA 6957
```

RESULT 5
ADL13909/c
ID ADL13909 standard; DNA; 129588 BP.

AC ADL13909;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #441.

KM de; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603-P.

PA (INCYTE) INCYTE GENOMICS INC.

PI Jones KA, Schaffer A;

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

PS Disclosure; SEQ ID NO 441; 297bp; English.

XX The invention relates to a method of determining susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain comprising identifying whether the individual has at least one polymorphism in a polynucleotide encoding at least one of the protein listed in the specification. The methods, composition and agent are useful for modulating the susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polynucleotide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences).

Sequence 129588 BP; 35710 A; 27530 C; 26424 G; 36620 T; 0 U; 3304 Other;

Query Match 7.0%; Score 351; DB 10; Length 129588;

Best Local Similarity 66.6%; Pred. No. 3e-48;

Matches 623; Conservative 0; Mismatches 275; Indels 37; Gaps 7;

QY 1 TATAGGCCAATGCTGAGGCTCAAGGCTGATCCAGCACTTGGAGGC---AGAGGA 57

DB 23050 TAAAGCTGTGAGTGGCTCACTCACTCAATCCAGTCTTAAGAGGCAAGAGGA 22991

QY 58 TCGCTTGAAGCTGAGGATTTGAGAGCAAGCTCACTGTAATGAATCACTGCTGTAC 117

DB 22990 TCACCTTGAAGCTGAGGATTTGAGAGCAAGCTCACTGTAATGAATCACTGCTGTAC 22931

QY 118 AATTAATTAAGAAATTTCCAGGATGCTGGGCTGACCCCAAGTCCAGCTAATTTGGGA 177

DB 22930 AAAAATAATTAAGTGTGGGCAAGGCTCACTCACTAATCTCACTTGGGA 22871

QY 178 GGCTGAGTGAAGGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 237

DB 22870 GGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 22811

QY 238 AGACCTGTGTCTAATTAATAAT---AATTAAGTGTGTGTGTGTGTGTGTGTGTGT 290

DB 22810 AATCTGATCTCTACTAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 22751

QY 291 GCTAGCTACTCGAAGACTGAGGAGGAGGATCAC-TGAGCCAGAGAGGCTGAGGCTGA 349

DB 22750 CCAGCTACTTGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 22691

QY 350 GTGAACAGTATCACCCAGCTGATTCAGGCTGAGAGAGAGAGAGAGGAGGAGGAGG 409

DB 22690 GTGAACAGTATCACCCAGCTGATTCAGGCTGAGAGAGAGAGAGGAGGAGGAGG 22631

QY 410 AAAAAAAAAAAAAAAAAATGCAAGAAAGACATCATTAATCTTGAGGAGGAGGAGG 469

DB 22630 AATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 22584

QY 470 TTTATGATGAATAATTCACAACTTTTGAAGAAATTAAGATTTCTGATTAATTAAT 529

DB 22583 TGTGTAGTCCCAATTCACAAAGGCTGAGGAGAGAGAAATCTTGAGAGGAGGAGG 22524

QY 530 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 589

DB 22523 AGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 22464

QY 590 CTCTCAGTTGTGAGTCAAACTTTAATGAGTCTTTGGCCGGGAGGAGGAGGAGGAGG 649

DB 22463 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22416

QY 650 AATCCAGCACTTTGGAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 709

DB 22415 AATCCAGCACTTTGGAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 22356

QY 710 CCTGCTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 769

Db 22355 CCTGGCTAACACGGTGAACCTGCTCTCTACTTAAATAATC-AAAAGTTAGCAGGCGTGG 22297
Qy 770 TCCAGACCCCTGTAGTCTCCAGCTGCTCAGAGGCTGAGGACGAGAAATGTTGAACCC 829
Db 22296 TGGCGGTGCTGTATGATCCAGTACTCGGAGGCTGAGACGAGAAATGGGTGAACCT 22237
Qy 830 GGGAGGCGGAGCTTGGAGTGAAGCCGAGATTGGCGCCACTGCACTCCAGCTTGGCGGACAT 889
Db 22236 GGGAGGCGGAGCTTGGAGTGAAGCCGAGATTGGCGCCACTGCACTCCAGCTTGGCGGACAG 22177
Qy 890 GCGAGACTGCTGTCTCAAAAAAAAAAAAAAAAAA 924
Db 22176 GCAAGACTTCATTCAAAAAAAAAAATTAATTA 22142

RESULT 7
ABQ8096/c
ID ABQ8096 standard; cDNA; 92638 BP.
XX
XX ABQ8096;
AC
XX 18-SBP-2002 (first entry)
XX
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 3.
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
osteoporosis; osteopathic; ss.
XX
XX Homo sapiens.
XX
XX MO200250301-A2.
XX
XX 27-JUN-2002.
XX
XX 18-DEC-2001; 2001MO-US048276.
XX
XX 18-DEC-2000; 2000US-0255882P.
XX
XX 24-APR-2001; 2001US-0285691P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX (PROC) PROCTER & GAMBLE CO.
XX
XX JI D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
DR MPI; 2002-557663/59.
XX
XX
XX Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
XX Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.
XX
XX The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring rate of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIP0
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;

Query Match 7.0%; Score 347.6; DB 6; Length 92638;
Best Local Similarity 65.9%; Pred. No. 1.1e-477;
Matches 618; Conservative 0; Mismatches 299; Indels 21; Gaps 7;
Qy 5 GAGCAATCTGTGGCTCAGCGGTATTCAGACATTTGGGAGGAG-----GAGATC 59
Db 90650 GCGAATGCAATGGGTGACACCTGTAATCCAGACCTTTGGGAGTCCAAAGGAGGTGATC 90591
Qy 60 GCTTAGCTCAGAAATTTGAGACAAAGCTTAACATAGTGAACCTCTGTGTACAA 119
Db 90590 ACTAAGGCGCAGAGATTTGAGACCGCTGGCCAAACAGTGAACCTGTCTACTTAA 90531
Qy 120 ATATATTAAGAAATTTCCAGGATGTGGCGGCGACCCCAAGTGCAGCTATTGGGAGG 179
Db 90530 AAAATACAAAATTTGGCTGAGATGTGGCTTCACACCTGTATCCAGACATTTGGAGG 90471
Qy 180 CTGAGGTAGAGAAATGCTTGAAGCCAGAGATTGAAGCAAGCTTAGCAACATATGAG 239
Db 90470 CCGAGGAGGCTGATCAGCTGAGTCAAGAGTTCAGAGACAGCCTGTCCAAAGTGTGA 90411
Qy 240 ACCCTGTGCTTATTAATAATTAATTAAGT-----GTTGTCTTGGCAGAGCTGCAGC 292
Db 90410 ACCCATCTCTAATTAATAATTAATAAATTTGCTGGGCAATGGCAACCTGTATTC 90351
Qy 293 TAGCTACTCGGAAGACTGAGGTGGGAGATCAC-TGAGCCGAGAGGCTGAGGCTGAGT 351
Db 90350 CAGCACTCGGAGGCTGAGGCTGAGGAGAGATGCTTGAACCCAGAGGAGAGGTTGCTGT 90291
Qy 352 GAACAGTATCACCCAGCTGATTCAGACCTGGAAGACAGAGGAGACCTGTTTCAA 411
Db 90290 GAGTAGAGATCGCGCATTCATTCAGACTAGTGAACA-AGAGGAAATCTGTCAAAA 90232
Qy 412 AAAAAAAAAAAAAAAAAATGCAAGAAAGATCTCTAACTTGAACCTGGACATTAATTT 471
Db 90231 AAAAAAAAAAACAACAACAACAACA-ACAAAACAACAACAACAACAACAACA 90173
Qy 472 TATGTAGTAAATTAACAATCTTTAGAGAGAAATTAACATTTCTGATTAATATATTA 531
Db 90172 AAAATTAATGCGGCGGAGTGTCTCAGCCTGTATCTAGCATGTGGAGCGGAGGT 90113
Qy 532 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 591
Db 90112 GGTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 90053
Qy 592 CTCAGTGTCACTCAAACTTTA--ATGCTCTTGGCGGCGGCTGAGGCTGAGCCTGT 649
Db 90052 CTCTACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 89993
Qy 650 AATCCAGCACTTTGGAGGCGGAGGCGGCTGATCAGAAAGTTAGAGATGAGACCAT 709
Db 89992 GGTCCAGACATTTGGAGGCGGAGGCGGCTGATCAGAGGTCAAGAGATGAGACCAT 89933
Qy 710 CCTGGCTAACACGGTGAACCTGCTCTCTAA-----ATATCAAAAATTAATAGCGGCT 765
Db 89932 CCTGGCTAATATGAGAAACATGCTCTCAAAAAAAAAAATTAATTAATTAATTAATTAAT 89873
Qy 766 GGGGTGCCAGAGCCGCTGATGCTCCAGCTGCTCAGAGGCTGAGGAGGAATGAGTGA 825
Db 89872 GGTGTGATGAGGACCTGTAAACCCAGCTGATTTGGAGGCTGAGGAGGAATGCTTGA 89813
Qy 826 ACCCGGAGGCGGAGCTTGCAGTGAAGCCGAGATTGGCGCACTGCACCTCAGCTGGCGGA 885
Db 89812 ACCCGGAGGCGGAGGTTGCAATGAGCCAAAGTCAACACACTGCACCTCGGCGTGGGCAA 89753
Qy 886 CAGTGGAGACTGTGTCTCAAAAAAAAAAAAAAAAAA 923
Db 89752 CAGAGTGAAGACTCAGTCTCAAAAAAAAAAATGATTAATTA 89715

RESULT 8
AD112698
ID AD112698 standard; DNA; 97955 BP.
XX

AC ADI12698;
 XX 22-APR-2004 (first entry)
 XX Human LPIN2 genomic DNA as given by ENSEMBL (Fig 3).
 DE de; type 2 diabetes; insulin resistance; human; lipin 2; LPIN2;
 XX linkage disequilibrium polymorphism; antidiabetic; gene therapy.
 KM Homo sapiens.
 OS MO2004001071-A2.
 PN 9
 XX 31-DEC-2003.
 PD 25-JUN-2003; 2003WO-GB002730.
 XX 25-JUN-2003; 2003WO-GB002730.
 XX 25-JUN-2002; 2002GB-00014682.
 PR (OXAG-) OXAGEN LTD.
 XX Pullen J, Holdstock J;
 PA WPI; 2004-082513/08.
 XX Determining whether an individual is predisposed to type 2 diabetes
 PT and/or insulin resistance, useful for treating and/or preventing such
 PT disease, comprises typing the LPIN2 gene region or LPIN2 protein of the
 PT individual.
 XX Disclosure; Fig 3; 152pp; English.
 PS This invention relates to a novel method for determining whether an
 CC individual is predisposed to type 2 diabetes and/or insulin resistance.
 CC Specifically, it comprises typing the human lipin 2 (LPIN2) gene in order
 CC to detect at least one of the four recognised single nucleotide
 CC polymorphisms (SNPs) known to be associated with type 2 diabetes and
 CC insulin resistance. The present invention further describes detecting
 CC linkage disequilibrium polymorphisms that indicate a susceptibility or
 CC genetic predisposition to these conditions. The method comprises
 CC contacting a test agent with an LPIN2 mutated polymucleotide or
 CC polypeptide and determining whether it is capable of binding and/or
 CC modulating the activity or expression of the molecule. Accordingly, these
 CC compositions exhibit antidiabetic activity and can be used to treat type
 CC 2 diabetes and/or insulin resistance using gene therapy. This
 CC polymucleotide sequence is a human LPIN2 genomic DNA sequence of the
 CC invention.
 XX Sequence 97955 BP; 25989 A; 19721 C; 20487 G; 31058 T; 0 U; 700 Other;
 SQ
 Query Match 6.9%; Score 346.6; DB 12; Length 97955;
 Best Local Similarity 65.7%; Pred. No. 1.6e-47;
 Matches 610; Conservative 0; Mismatches 289; Indels 30; Gaps 6;

QY 329 GCCCAGAGAGCTGAGGCTGCACTGAACTGATCAACCCAGCTGATTCAGAGCTGGAAGA 388
 DB 15329 ATCCGGAGGTGAGGTGCAATGAGCCGAGATCTGTCACCGCACTCCAGCTTGAGTGA 15388
 QY 389 CAGAGGAGAGACCTGTTCCAAAAAATGCAAG-----AAAGA 441
 DB 15389 GGGAGTGAGACTCTGTTCCAAAAAATGCAAG-----AAAGA 15448
 QY 442 CATCATAACTTACCTGGGACATTAATTTATGATGAAATTCACAACTTTTAAAGAA 501
 DB 15449 AATGTACATATTATGATGATGTTGATTCGCACTTATGCAAGATTTACATATAC 15508
 QY 502 GAAATTAAGATTTTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 561
 DB 15509 AATATATAGCAAACTTGTGCTCCATGCAAGGCTTTGTGAACCAAAATGATCTTTATTT 15568
 QY 562 AATATTTCTGAGAACTAGCTTCTCACTCTCACTGTCAGTGTCACTCAAACTTTAATGTC- 620
 DB 15569 GAGATGCTGGAGATGAGGCTCTGTTATTTCTGGGATATATATATCTCAATAGATAGATA 15628
 QY 621 -----TTTGGCCGGGTGCGGTGCTCAGCCTGTAATCCAGACTTTTGGAGCCGAG 675
 DB 15629 AAGATATTGCTGAGGCTGCTGCTCAGCCTGTAATCCAGACTTTTGGAGCCGAG 15688
 QY 676 CCGGTGATCACAAGGTTAGAGATGAGACCACTCTGCTAAACAGCTGAAACCTCTCTC 735
 DB 15689 TGGGTGATCAGAGGTCAGAGATGAGACCACTCTGCTAAACAGCTGAAACCTCTCTC 15748
 QY 736 TCTACTAAAAATACAAAAAATATAGCCGGGTGCGGTGCTCAGCCTGTAATCCAGACTTC 795
 DB 15749 TCTACTAAAAATACAAAAAATATAGCCGGGTGCGGTGCTCAGCCTGTAATCCAGACTTC 15808
 QY 796 TCAGAGGCTGAGGAGAGAAATGTTGTAACCCGGAGGCGGAGCTTGCAGTGAACCGCA 855
 DB 15809 TCGGAGGCTGAGGAGAGAAATGTTGTAACCCGGAGGCGGAGCTTGCAGTGAACCGCA 15868
 QY 856 GATTGGCCCATCTGCACTCTCAGCTGGGCGACAGTGGAGACTCTGTTCTAAAAAATGAA 915
 DB 15869 GATTGGCCCATCTGCACTCTCAGCTGGGCGACAGTGGAGACTCTGTTCTAAAAAATGAA 15928
 QY 916 AAAAAAAGTGAATGCTTGAAGCA 944
 DB 15929 AAAAAAAGTGAATGCTTGAAGCA 15957

RESULT 9
 AB075562
 ID AB075562 standard; DNA; 188888 BP.
 AC AB075562;
 XX 11-NOV-2002 (first entry)
 DT Human related CYP 27C1 clone RP11-30F3 SEQ ID NO:21.
 XX Cloning; characterisation; human; cytochrome P450; CYP 27C1; cytostatic;
 XX chymotrypsin; antidiabetic; antihypertensive; tuberculostatic; osteoporosis;
 XX dermatologic; antidiabetic; gene therapy; vaccine; Vitamin D; diabetes;
 XX vitamin D metabolite deficiency; hyperparathyroidism; hypoparathyroidism;
 XX medullary carcinoma; psoriasis; sarcoidosis; tuberculosis; osteomalacia;
 XX chronic renal disease; vitamin D dependent rickets; anticonvulsant;
 XX fibrogenesis imperfecta ossium; osteitis fibrosa cystica; osteoporosis;
 XX osteopenia; osteosclerosis; renal osteodystrophy; rickets; steatorrhoea;
 XX glucocorticoid antagonism; idiopathic hypercalcaemia; tropical sprue;
 XX malabsorption syndrome; cholesterol steroid; lipid metabolic disorder;
 XX gene; de.
 OS Homo sapiens.
 XX MO200264765-A2.
 PN 22-AUG-2002.
 XX

Sequence 59588 BP; 15931 A; 13628 C; 13537 G; 16492 T; 0 U; 0 Other;
Query Match 6.9%; Score 345.2; DB 10; Length 59588;
Best Local Similarity 68.9%; Pred. No. 2.6e-47;
Matches 554; Conservative 0; Mismatches 203; Indels 47; Gaps 4;
QY 136 CCAGGCTAGTGGCCGCGACCCCGAGTGGCAGTATTTGGAGGCTGAGTGAAGT 195
DB 46072 CCAGGCTATATGCTCTCATGCTCTTAATCCCAAGTTGGAGGCTCAAGTGGAGGAT 46013
QY 196 GCTTGAAGCCAGAGTGAAGAACAAGCTTAGCAATAGTGAAGCTGTGTCTA-TAA 254
DB 46012 GCTTGAAGCCAGAGTGAAGAACAAGCTTAGCAATAGTGAAGCTGTGTCTA-TAA 45953
QY 255 AAAATTAATTAGTGTGTCTTGGCAAGAGCTGAG-CTAGCTACTGGAAGCTGAG 312
DB 45952 AAAATTAATTAGTGTGTCTTGGCAAGAGCTGAG-CTAGCTACTGGAAGCTGAG 45893
QY 313 GTGGAGAGATC-TCAGCCAGAGGCTGAGGCTGCAATGAACTGATCAACCCAGCTG 371
DB 45892 GTGGAGAGATC-TCAGCCAGAGGCTGAGGCTGCAATGAACTGATCAACCCAGCTG 45833
QY 372 GATTCAGGCTGGAAGACAGAGGAGACCTGTCTTCAAAAAAATTTTCAAAAAA 431
DB 45832 TACTTCAGGCTGGAAGACAGAGGAGACCTGTCTTCAAAAAAATTTTCAAAAAA 45773
QY 432 CAAGAAAAGACATCAATTAATCTGAGGACATTAATTTATGTGATGAAATTCACAT 491
DB 45772 AGACAAATATGTGCTGCTGCTGCAATTCAGACATTTGAGAGCCAGAGGAGATC 45713
QY 492 CTTTATGAGAAATTAATGATTTTGATTAATTAATTAATTAATTAATTAATTC 551
DB 45712 TCTTATGAGAAATTAATGATTTTGATTAATTAATTAATTAATTAATTAATTC 45653
QY 552 AAATGAAATTAATTAATTTGAGAAATTAATCTTCTGATCTGATGATGATCAAACT 611
DB 45652 AAATGAAATTAATTAATTTGAGAAATTAATCTTCTGATCTGATGATGATCAAACT 45628
QY 612 TTAATGTCTTTGCGCGGCTGCGGCTGCTGAGGCTGTAATCCAGCACTTTGGAGGC 671
DB 45627 TTAATGTCTTTGCGCGGCTGCGGCTGCTGAGGCTGTAATCCAGCACTTTGGAGGC 45576
QY 672 GAGGCGGCTGATCAAGGTTAGAGATTCAGACATCTCTGCTTAACCGGTGAACCT 731
DB 45575 GAGGCGGCTGATCAAGGTTAGAGATTCAGACATCTCTGCTTAACCGGTGAACCT 45516
QY 732 CGTCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 791
DB 45515 TGTCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 45456
QY 792 CTGCTCAGAGGCTGAGGAGAGAGATGTGTAACCCGGAAGGCGGAGCTTGCAGTGA 851
DB 45455 CTGCTCAGAGGCTGAGGAGAGAGATGTGTAACCCGGAAGGCGGAGCTTGCAGTGA 45396
QY 852 CCGAATATGCGGCTGAGGAGAGAGATGTGTAACCCGGAAGGCGGAGCTTGCAGTGA 911
DB 45395 CCGAATATGCGGCTGAGGAGAGAGATGTGTAACCCGGAAGGCGGAGCTTGCAGTGA 45336
QY 912 AAAAAAAAAAAGTTGAATGATC 935
DB 45335 TAAAAAAAAAATAATTAAGCC 45312

RESULT 11
AD112699
ID AD112699 standard; DNA; 95103 BP.
XX AD112699;
AC
XX 22-APR-2004 (first entry)
DT
XX Human LPIN2 genomic DNA as given by Genome Browser (Fig 4).
DE
XX

KM de; type 2 diabetes; insulin resistance; human; lipin 2; LPIN2;
KW linkage disequilibrium polymorphism; antidiabetic; gene therapy.
XX
OS Homo sapiens.
XX
PN MO2004001071-A2.
XX
PD 31-DEC-2003.
XX
PF 25-JUN-2003; 2003WO-GB002730.
XX
PR 25-JUN-2002; 2002GB-00014682.
XX
PA (OXAG-) OXAGEN LTD.
PI Pullen J, Holdstock J;
XX WPI; 2004-082513/08.
DR
XX
XX
PT Determining whether an individual is predisposed to type 2 diabetes
PT and/or insulin resistance, useful for treating and/or preventing such
PT disease, comprises typing the LPIN2 gene region or LPIN2 protein of the
PT individual.
XX
PS Disclosure; Fig 4; 152pp; English.
XX
XX This invention relates to a novel method for determining whether an
CC individual is predisposed to type 2 diabetes and/or insulin resistance.
CC Specifically, it comprises typing the human lipin 2 (LPIN2) gene in order
CC to detect at least one of the four recognized single nucleotide
CC polymorphisms (SNPs) known to be associated with type 2 diabetes and
CC insulin resistance. The present invention further describes detecting
CC linkage disequilibrium polymorphisms that indicate a susceptibility or
CC genetic predisposition to these conditions. The method comprises
CC contacting a test agent with an LPIN2 mutated polynucleotide or
CC polypeptide and determining whether it is capable of binding and/or
CC modulating the activity or expression of the molecule. Accordingly, these
CC compositions exhibit antidiabetic activity and can be used to treat type
CC 2 diabetes and/or insulin resistance using gene therapy. This
CC polynucleotide sequence is a human LPIN2 genomic DNA sequence of the
CC invention.
XX
SQ Sequence 95103 BP; 25304 A; 19358 C; 19984 G; 29657 T; 0 U; 800 Other;
Query Match 6.9%; Score 344; DB 12; Length 95103;
Best Local Similarity 66.0%; Pred. No. 4.1e-47;
Matches 602; Conservative 0; Mismatches 280; Indels 30; Gaps 6;
QY 46 GAGGCAAGAGATCGCTGAGCTCAGAAATGGAGACAAAGCTTACATAGTGAAC 105
DB 38917 GAGGAGGCTGATCAATGAGGCTCAGAGATTCAGACAGCTGCGCAACAGATTAAC 38976
QY 106 CTCTCTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 156
DB 38977 CCATCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39036
QY 157 CCGAGTGCAGCTATTTGGAGGCTGAGAGTGAAGAAATGTTGAAGCAGAGATTGAAG 216
DB 39037 TGTATCCAGTACTTTGGGACACCAAGGCGGCGGATCACTGAGTCAAGAGATTGA 39096
QY 217 ACAAGCTTAGGCAATAGTGAACCTGTGTCTTAATAAATAAATAAATAAATAAATAAATA 271
DB 39097 ACAAGCTTAGGCAATAGTGAACCTGTGTCTTAATAAATAAATAAATAAATAAATAAATA 39156
QY 272 GTCTTGACACAGGCTGCA-GCTAGCTACTGGAAGCTGAGGAGGAGATCA-TGA 328
DB 39157 GTGGGGGTGGTACTGTAACTTCTAGTCTGAGAGCTGAGAGAGAAATTTCTTGA 39216
QY 329 GCCAGAGAGCTGAGGCTGAGTGAACAGTGAATCAACAGCTGATTCAGGCTGGAAGA 388
DB 39217 ATCCGAGAGGTGAAGTTCATAGAGCGGAGATGCTGACCGCACTCAAGCTGGTGA 39276
QY 389 CAGAGGAGACCTGTCTTCAAAAAAATTTTCAAG-AAAAA 441

PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 01-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-02559678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 36006; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the

CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX Sequence 9979 BP; 2496 A; 2022 C; 2532 G; 2929 T; 0 U; 0 Other;
 Query Match 6.9%; Score 343; DB 4; Length 9979;
 Best Local Similarity 67.2%; Pred. No. 5.5e-47;
 Matches 627; Conservative 0; Mismatches 280; Indels 26; Gaps 9;
 QY 13 CTGTGGCTACGCGGTGTATTCACACCTTTGGAGC-----AGAGGATCGTTGAG 66
 DB 9080 CAGTGGCTCACACTGTATATCCACACACTTGGAGGCGCAAGTGGTGTATCACTTGAG 9021
 QY 67 CTCAGGAATTGGAGCAAGCCCTACGTAACATAGTGAACCTGTCTGTACAAATATA 126
 DB 9020 GTTCAGAGATTAGACAGCCCTGGCCACAGGGTGAACCCCATCTCTAATAAATAGA 8961
 QY 127 AAGAAATTTCCAGGATGGTGGCGTGACCCCCAGTGCAGCTAATTTGGAGGCTGAGGT 186
 DB 8960 AGAATTAGCCACGCAATGTGCTCACACCTTAATCCAGACCTTTGGAGGCGGAGGC 8901
 QY 187 AGAGGAATGCTTGAAGCCAGAGATTGAAGACAAAGCCTTAGCAATAGTGAACCTTGT 246
 DB 8900 AGGTG--AATCAGAGGTGAGAGATCAAGACATCTGTGATTAACGGTGAACCCCGT 8843
 QY 247 GTCTATN-----AAAAATATTAGCTGTGTCTTGGACACAGGCTGACG--CTAGCT 297
 DB 8842 TCTACTAAATAACAAAAAATTTAGCCGGGATGTGAACAGCCCTGTATCTCCAGCT 8783
 QY 298 ACTCGGAAGACTGAGTGGAGAGATCAC--TGAGCCAGAGGCTGAGGCTGAGTGAACA 356
 DB 8782 ACTTAGAGGCTTAAGGACGAGAAATCACTTAACCTGAGAGGACAGAGCTGAGGAGGC 8723
 QY 357 GTGATCACCCAGCTGGAATTCAGCCTGGAAGACAGAGGAGACCTGTGTTCCAAAAA 416
 DB 8722 GAGATCACGCCCACTACACTCCAGCTGGGCGACAGAGGAGACTGTGCACCAAAAAA 8663
 QY 417 AAAAAAAAAAATGCAAGAAAGACATCATTAACCTGACGCG---ACATACTTT 472
 DB 8662 AAAAAACAGAAAGCTGTATTAGTATTTGAACCAAGTTCACATCATATATAATGACA 8603
 QY 473 ATGTATGAATTCACAAATCTTTTGAAGAAATAGCATTTCTGATAAATGATATTA 532
 DB 8602 AAGCGTTACTTCACTGATGATCAAGTCAATCTTCAATTATGATGAGCCACACA 8543
 QY 533 ATTATATTATTATTAATTAATGAAATTAATATTGAGAACTAGCTTCTCACTTC 592
 DB 8542 AAGAAACAAAAGGGGATCATTTGACAAATGTA--TGCCAGACATTCACATTTGTAT 8484
 QY 593 TGAATTGATGACAAAATCTTAATAGTCTTTGGCCGGGTGCGGTGCTCACCCCTGTA 652
 DB 8483 TCAATTGTCTCTTAAGAAATTTCTCTTCGTGGTGGGTGCGGTGCTCACCCCTGTA 8424
 QY 653 CCCAGACATTTGGAGGCGGAGGCGGGTGAATCAACAGTTAGAGATGAGACCATCT 712
 DB 8423 CCCAGACATTTGGAGGCGGCGGCGGGTGAATCAAGAGTCAAGAGATCAACACATCT 8364
 QY 713 GGCTAACACGCTGAACCTGCTCTCTACTAATAAATACAAAAAATTAGCCGGGTGCGTGC 772
 DB 8363 GGCTAATACGGTGCAACACATCTCTACTAATAAATAC--AAAAACAAAATTAGCCGGGTGC 8305
 QY 773 CAGAGCGCTGATGCTCCAGCTGCTTACAGAGGCTGAGGACAGAAATGCTGAA--CCGC 830
 DB 8304 TGGGCGCTGATGCTCCAGCTACTCGGAGGCTGAGGACAGAAATGCTGAAATCTCGC 8245
 QY 831 GGAGCGGAGCTTGCAGTGAAGCCGATTTGGCCACTGCACTCCAGCTGGGCGACAGT 890
 DB 8244 GGAGGTGAGCTTGCAGGAGCTGAGATCTGCACTCCAGCTGGGCGACAG 8185

Yr	891	CGAGACTGTGTCACAAAAAAAAAAAAAAAAAAAA	923
Db	8184	CAGACTTGTCTCAAAAAAAAAAAAAAAAAAAAAA	8152
RESULT 13			
ADP75180			
ID	ADP75180	standard; DNA; 304905 BP.	
AC	ADP75180;		
XX			
XX	12-AUG-2004	(first entry)	
XX			
XX	Human Endophilin 2 gene.		
XX			
XX	Human; chromosome 19; ds; gene; ADAM9; Endophilin 1; Endophilin 2; NR32;		
XX	ADAM752; a disintegrin and metalloprotease; neuroregulin 2; SNP;		
KW	single nucleotide polymorphism;		
KW	a disintegrin and metalloprotease with thrombospondin type1 motif 2;		
KW	asthma; atopy; obesity; inflammatory bowel disease; respiratory disorder.		
XX			
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
XX	variation	replace(276365,T)	
XX		/*tag= a	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(276366,G)	
XX		/*tag= b	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(278194, .278197,CT)	
XX		/*tag= c	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(278474,A)	
XX		/*tag= d	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(278990,C)	
XX		/*tag= e	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(279072,G)	
XX		/*tag= f	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(280049,T)	
XX		/*tag= g	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(280107,A)	
XX		/*tag= h	
XX	variation	/standard_name= "Single nucleotide polymorphism"	
XX		replace(280124,T)	
XX		/*tag= i	
XX		/standard_name= "Single nucleotide polymorphism"	
XX			
XX	WO2003031594-A2.		
XX			
XX	17-APR-2003.		
XX			
XX	11-OCT-2002; 2002WO-US032700.		
XX			
XX	11-OCT-2001; 2001US-0328424P.		
XX			
PA	(GENO-) GENOME THERAPEUTICS CORP.		
XX			
PI	Keith T, Little RD, Van Berdeghen P, Dupuis J, Del Maestro RG;		
PI	Allen K;		
XX			
DR	WPI; 2003-381712/36.		
XX			
PT	New isolated nucleic acid or alternate splice variant, useful for		
PT	diagnosing and treating a disintegrin and metalloprotease (ADAM) or		
PT	interactor gene-associated disorder, e.g. asthma, atopy, obesity or		
PT	inflammatory bowel disease.		
XX			

Query Match	Best Local Similarity	6.8%;	Score 341.2;	DB 11;	Length 304905;
Matches 625;	Conservative 0;	Mismatches 253;	Indels 52;	Gaps 8;	
QY	10	ATGCGTGGGCGACGCGGTATTCGCCGACCTTGG-----GAGCGAGAGAGATGCTT	63		
Db	24379	ATGCGATGGCTCAAGCTCTGTAATCTAGACCTTTGGGAGTGTGAGCGACGATCACT	24438		
QY	64	GAGCTCAGGAATGAGACAAAGCTTACATAGTAAACCTCTGCTGTATCAAAATTA	123		
Db	24439	GAAGTGGGAGATTGAGACTAGCCTGCGCAACATGGTGAACC--CCATCTCCATTA	24497		
QY	124	TAAAGATTTTCCAGGATGTGTGGGTGACCCCGGATGTCGACGATTTTGGAGGCTGA	183		
Db	24498	TACAAAAT-----ATGTGGCCCAACCTGTATCCGACGACTTTGGAGGAGCA	24548		
QY	184	GGTAGAGAAATGCTTGAAGCAGAGATTGAACAAAGCTTGAACATATGTAGACCC	243		
Db	24549	GCGGGGTGATCACTGTAGTCACGAGATTCACAGACCGCTGCGCAACATGGCAAAACC	24608		
QY	244	TGTGTCTTAAAAAT-----AATTGCTGTGTCTTGGCACAGGCGTGC--GCTAG	295		
Db	24609	CGTCTCTACTAAATAACAAAAATATGCTGGGCACTGTGTGTGGCACCTGTATATCCAG	24666		
QY	296	CTACTCGAAGACATGAGGTGGAGATCAC--TGAGCCCAAGAGGCTGAGGCTCACTGA	354		

PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242232P.
PR 01-NOV-2000; 2000US-0244177P.
PR 08-NOV-2000; 2000US-0246174P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 06-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-541565/60.
XX
XX Nucleic acid encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
PS Disclosure; SEQ ID NO 6912; 1701bp + Sequence Listing; English.
XX

CC The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 8288 BP; 2622 A; 1567 C; 1526 G; 2573 T; 0 U; 0 Other;

Query Match 6.8%; Score 341; DB 5; Length 8288;
Best Local Similarity 68.3%; Pred. No. 1.1e-46;
Matches 566; Conservative 0; Mismatches 245; Indels 18; Gaps 6;
QY 115 TACAATATATAAGATTTTCCAGCATGTGCGTGCACCCCACTGCGCACTTATTTG 174
DB 5669 TACCATATTAAAGTAAGACTTGCGACGCTGGACACACTTAACTTAGACCTTTG 5610
QY 175 GGAGGCTAGGTAGAGGAAATGCTTGAAGCCAGAGATTGAAGACAAAGCTTGGCAACTA 234
DB 5609 GGAGGCCGAAAGTATGATCATATGACAGCAAAATTGAGACACAGCTGCAACATG 5550
QY 225 GGGAGACCCCTGCTATATAAAAT-----AATTAGCTGTGTCTTGGCAGAGCTTGC 289
DB 5549 GTAAACCTCATCTTTACTTAAATAATACAAAATTAGCCGGCGCTGTATCCGACACCAT 5490
QY 220 AG---CTAGCTACTCGAAGACTGAGTGGAGGATCAC-TGAGCCAGAGGCTGAGGC 345
DB 5489 AGTCCAGCTAATCTCGGAGGCTAAGACAAAGAAATCATCTTGAACCTTGGGACAGAGGT 5430
QY 346 TCGAGTGAACAGTATCACCCAGCTGATTCACGCTGGAAGACAGAGGAGACCTGTT 405
DB 5429 TCGAGTAACTGAGATCACTCGCTGCACTCAGCTTG--AGTGAAGTGAAGACTGTGTC 5372
QY 406 TCCAAAAAATAAATAAATAAATGCAAGAAAGACATCATTAATTCAGCTGGACAT 465
DB 5371 TCAACAAACAAACAAACAAACAAACAAATTTGCAATTTGAATTTGACCTCAATTT 5312
QY 466 AACTTTATGATGATAATTCACAATCTTTAGAGAAATTAGCAATTTTCATTAATG 525
DB 5311 GAGCAGTTTTGTGTGTGTGTAATTAAGTAAGTAAGCAAGTCCACAAATTAATTTGTGA 5252
QY 526 TATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 585
DB 5251 TACTGACACAGTTCACACAGTTCACAGTTCACAGTTCACAGTTCACAGTTCACAGTTC 5192
QY 586 CACTCTCAGTTGTGAGTCAAACTTTAATGCTTTTGGCCGGGTGCGGTGCTCAGC 645
DB 5191 AAGAGCTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5136
QY 646 CTGTAAATCCAGCACTTTGGAGGCGGAGGCGGAGTCAACAAAGTTAGAGATCGAGA 705
DB 5135 TTGTATATCCACACTTTGGAGGCGGAGGCGGAGTCAACAAAGTTAGAGATCGAGA 5076
QY 706 CCATCTGCTGAACACCGGTAACCTCGCTACTAATAAATATAC---AAAAATTAGCCG 762
DB 5075 CCATCTGCTGAACACCGGTAACCTCGCTACTAATAAATATATAGCCG 5016
QY 763 GGTGCGTGCAGACGCTGTAGTCCAGCTCTCAGAGGCTGAGGACGAGAAATGTG 822
DB 5015 GCGGTATAGCGCGCGCTGTAGTCCAGCTCTCAGAGGCTGAGGACGAGAAATGTGCG 4956
QY 823 TGAACCCGGAGGCGAGCTTGCAATGAGCCAGATTTGCGCCACTGCACTCCAGCTGCG 882

Db 4955 TGAACCGGGAGGCGGAGCTTGACAGTACCGGAGATTGCCCACTGCACTCCGCTGGG 4896
Qy 883 CGACAGTGGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAGTTGAAT 931
Db 4895 CCACAGAGGAGACTCGTCTCAAAAAAAAAAAAAAAAAAGAAAT 4847

RESULT 15
AAK85745/c
ID AAK85745 standard; DNA; 4515 BP.
XX AAK85745;
AC
XX 07-NOV-2001 (first entry)
DT
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40557.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-020515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225157P.
XX 14-AUG-2000; 2000US-0225158P.
XX 14-AUG-2000; 2000US-0225159P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0228927P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR

PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251566P.
 PR 08-DEC-2000; 2000US-0251688P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 08-DEC-2000; 2000US-0251909P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI, 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 40557; 3071bp + Sequence Listing; English.

XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX Sequence 4515 BP; 1075 A; 1192 C; 1099 G; 1145 T; 0 U; 4 Other;

Query Match 6.8%; Score 338.6; DB 4; Length 4515;
 Best Local Similarity 67.1%; Pred. No. 2.8e-46;
 Matches 634; Conservative 0; Mismatches 239; Indels 72; Gaps 8;

QY 15 GTGGCTACGCGGTATTCCTCCAGCACTTTGGG-----AGGCAAGAGATCGCTTGAGC 67
 DB 941 GTGGCTATCCCTGTATCCAGCACTTTGGGAGGCCAAGGCGGGAGATCACTGAGC 882
 QY 68 TCAGGAATGGAGCAAGCCTACGTAACATAGTGAACCTCTGTCTGTACAAATAATAA 127
 DB 881 CCAAGATTCAAGACAGCCTGGGCAACATAGTGAACCTCTGTCTCAAAAAAAAAAAAA 822
 QY 128 AGAATTTTCA-----GGCATGTGGCGGTGACCCCGCAGTCCAGCTAATT 173
 DB 821 AAAAAAAAAAAGAGAGGTGGGTGGGTGCTCACCCCTGTATCCAGCACTT 762
 QY 174 GGGAGGCTGAGGTAGAGGATGTTGAAGCCAGGAGTTGAAGCAAGCCTTAGCAACAT 233
 DB 761 GGGAGGCGCGAGGTGGTGA--TCACGAGGTCAAGGATTGAGCAATCCAGGCTAACAG 704
 QY 234 AGTGAACCCCTGTGTCT-----ATAAAAAAAAAATTAAGTGTGTGTGGACA 282

DB 703 GGTGAACCCCATCTCTACTAATAAATACAAAAAATATGACAGGCGTGTGTGG 644
 QY 283 GGCCTGCAG--CTAGCTACTCGAAAGACTGAGGTGGAGATCA--TAGCCCAAGAGC 339
 DB 643 CACCTGTAGTCCCACTACTTGGAGGCTGAGGCGAGAGAAATGGCGTGAACCCAGAGAGC 584
 QY 340 TAGGCTGAGTGAACATGATCAACCAGCTGATTCAGCTGGAGAACAGAGGAGAC 399
 DB 583 AAGGTTACAGTGAACCGAGATAGCGCACTACCTCCAGCTGGGCAAGAGCAAGAC 524
 QY 400 CCTGTTCCAAAAAATAAAAAAAAAAATGCAAGAAAGACATCATAACTTGACCTG 459
 DB 523 TCTGCTTAATAAATAAATAAAGAGAAAAAATAAATAAATAAATAAAGACGCGC 464
 QY 460 GGAATTAATTTATGTATGATGAATTCACATCTTTTGAAGAAATTAATTCAT 519
 DB 463 GGGCGGTGCTACCGCTGTATCCAGCACTTGGAGGCTGA----- 420
 QY 520 AAAAAATGATTAATAATTAATTAATAATTAATTAATTAATTAATTAATTAATTA 579
 DB 419 -----TGGGGTGAATCAAGGTCAAGATGAGATGACATCTTGGCTAACAGGT 369
 QY 580 GCTTCTCACTCTCTCAGTTGTCAATCAAACTTTAATGATGATGCGCGGTGCGTGGC 639
 DB 368 GAAACCCCGCTCTATATAAATAAATAA-----AATTGGCCGGGCGGTGGC 319
 QY 640 TCACGCTGTATTCOAGCACTTTGGAGGCCGAGCGGTGATCAAGATTAGAGA 699
 DB 318 TCACGCTGTATTCOAGCACTTTGGAGGCCGAGCGGTGATCAAGATTAGAGA 259
 QY 700 TCGAGCAATCTGGCTAACAGAGTGAACCTGCTCTACTAATAAATAAATAAATTAG 759
 DB 258 TCGAGCAATCTGGCTAACAGAGTGAACCTGCTCTACTAATAAATAAATAAATTAG 199
 QY 760 CCGGTTGCGGTCCAGACGCTGTAGTCCAGCTGCTGAGAGGCTGAGGCAAGATG 819
 DB 198 CCGGTTGCGGTCCAGACGCTGTAGTCCAGCTGCTGAGAGGCTGAGGCAAGATG 139
 QY 820 GTGTGAACCCGGAGAGCGAGCTTGCATGAGCCAGATTTGGCCACTGCCAGCT 879
 DB 138 GGTGAACCCGGAGAGCGAGCTTGCATGAGCCAGATTTGGCCACTGCCAGCT 79
 QY 880 GGGCGAGGTGAGACCTGTCTCAAAAAAAAAAAAAAAAAAAAA 924
 DB 78 GGGCGAGAGAGGACCTCTCTCAAAAAAAAAAAAAAAAAAAAA 34

Search completed: December 26, 2004, 18:48:42
 Job time : 1456.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:39:26 ; Search time 16925.5 Seconds
(without alignment)
17325.567 Million cell updates/sec

Title: US-10-786-065-3_COPY_20000_26200

Perfect score: 6201

Sequence: 1 caccatgggcccaggaattc.....aagagtaatgaagatctg 6201

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6201	100.0	53332	6	AR265351	AR265351 Sequence
2	6201	100.0	53332	6	AX571875	AX571875 Sequence
3	6129.2	98.8	112748	9	AC007242	AC007242 Homo sapi
4	419.8	6.8	145212	2	AC026170	AC026170 Homo sapi
5	419.8	6.8	185321	2	AC026153	AC026153 Homo sapi
6	412.8	6.7	2001	9	HS0808294	HS0808294 Homo sapi
7	412.8	6.7	2002	9	HS0807657	HS0807657 Homo sapi
8	402.6	6.5	1846	9	BC008944	BC008944 Homo sapi
9	402.6	6.5	1870	9	BC008944	BC008944 Homo sapi
10	402.6	6.5	1930	6	AR000150	AR000150 Sequence
11	402.6	6.5	1982	9	HS039360	HS039360 Homo sapien
12	402.6	6.5	2135	9	HS039361	HS039361 Homo sapien
13	402.6	6.5	3802	6	CQ413441	CQ413441 Sequence
14	387.2	6.2	229392	2	AC148906	AC148906 Ocolemur
15	386.6	6.2	219125	9	AC008736	AC008736 Homo sapi
16	386.6	6.2	216441	2	AC027340	AC027340 Homo sapi
17	381.8	6.2	36171	9	HS121988	HS121988 Human DNA
18	381.4	6.2	119582	2	AC022110	AC022110 Homo sapi
19	380.8	6.1	145151	9	AL645949	AL645949 Human DNA

C	20	378	6.1	181230	2	AC149856	AC149856 Papio anu
C	21	376.6	6.1	129564	9	AC087302	AC087302 Homo sapi
C	22	376.6	6.1	188800	2	AC015846	AC015846 Homo sapi
C	23	375	6.0	106921	9	AC004263	AC004263 Homo sapi
C	24	375	6.0	165496	9	AC104461	AC104461 Homo sapi
C	25	375	6.0	176861	9	AC097065	AC097065 Homo sapi
C	26	375	6.0	189729	2	AC073930	AC073930 Homo sapi
C	27	375	6.0	220384	2	AC078926	AC078926 Homo sapi
C	28	374.4	6.0	94161	9	AC020937	AC020937 Homo sapi
C	29	373	6.0	182245	2	AC149842	AC149842 Papio anu
C	30	372.4	6.0	149001	9	AC107909	AC107909 Homo sapi
C	31	372.4	6.0	153411	9	AC104986	AC104986 Homo sapi
C	32	372.4	6.0	180236	9	AP004294	AP004294 Homo sapi
C	33	372.2	6.0	197298	2	AC150024	AC150024 Papio anu
C	34	370	6.0	159692	9	AC016727	AC016727 Homo sapi
C	35	369.6	6.0	172850	9	AC009269	AC009269 Homo sapi
C	36	369.6	6.0	183890	9	AC090868	AC090868 Homo sapi
C	37	368.8	5.9	69110	9	AC069437	AC069437 Homo sapi
C	38	368.4	5.9	53595	9	AL607063	AL607063 Human DNA
C	39	367.8	5.9	157381	9	AL360267	AL360267 Human DNA
C	40	367.6	5.9	80206	2	AC138693	AC138693 Homo sapi
C	41	367.6	5.9	178455	2	AC146015	AC146015 Pan trogl
C	42	366	5.9	145706	9	AC092612	AC092612 Homo sapi
C	43	365.8	5.9	3335	9	HS049278	HS049278 Homo sapien
C	44	365.2	5.9	86453	9	HS1049G16	HS1049G16 Human DNA
C	45	365.2	5.9	149817	2	AC069440	AC069440 Homo sapi

ALIGNMENTS

RESULT 1	AR265351	Sequence 3 from patent US 6492154.	DNA	linear	PAT 10-APR-2003
LOCUS	AR265351	53332 bp			
DEFINITION	AR265351				
ACCESSION	AR265351.1	GI:29693854			
VERSION	AR265351.1	GI:29693854			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 53332)				
TITLE	Yan, C., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.				
JOURNAL	Isolated human kinase proteins, and uses thereof				
FEATURES	Patent: US 6492154-A 3 10-DEC-2002;				
SOURCE	Location/Qualifiers				
	1..53332				
	/organism="unknown"				
	/mol_type="genomic DNA"				

ORIGIN

Query Match	100.0%	Score 6201;	DB 6;	Length 53332;
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 6201;	Conservative	0;		Gaps 0;
QY	1	CACCATGGGCTCAGGCAATTTATGAAGCAATATACAGCTTAAATAGATGTGAC	60	
DB	20000	CACCATGGGCTCAGGCAATTTATGAAGCAATATACAGCTTAAATAGATGTGAC	20059	
QY	61	CTAAATACCCAGAGAACTCCCTTTGTAGATTGTGAACAAAATTAATAGTAGAG	120	
DB	20060	CTAAATACCCAGAGAACTCCCTTTGTAGATTGTGAACAAAATTAATAGTAGAG	20119	
QY	121	TTAATAGTTCTAATGAATGGTGAACCAAGGCAATATCAGGCTAGCAAAATGCGCA	180	
DB	20120	TTAATAGTTCTAATGAATGGTGAACCAAGGCAATATCAGGCTAGCAAAATGCGCA	20179	
QY	181	ATTCAATATATCATCAAGTTATCCTTCAAGAGCTTCAGGCTTAATGATGTCTAAAGAA	240	
DB	20180	ATTCAATATATCATCAAGTTATCCTTCAAGAGCTTCAGGCTTAATGATGTCTAAAGAA	20239	
QY	241	ATGTGAAGCGCCCTCAGGCAATCTGAAGGACAGTGTACAGCAATTGATCAAAAAAGAAAA	300	

20240 ATGTAAAGCCCTTACGCCATCTGAAGGACAGTGTACAGCAATGATCAAAAAAGAAAA 20299
QY |||||
301 CCACAGGCCCTTCCCTTCCCATATCTTGATGTAGAGAGCTTCTCAATTTTCCATAGTAGT 360
Db |||||
20300 CCACAGGCCCTTCCCTTCCCATATCTTGATGTAGAGAGCTTCTCAATTTTCCATAGTAGT 20359
QY |||||
361 AAATTTTCTAGATACAGCTTTGTAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTAAG 420
Db |||||
20360 AAATTTTCTAGATACAGCTTTGTAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTAAG 20419
QY |||||
421 GAAATTTATCTTAAGATACGTAAATGATTAATTTTGTACATTTGGAAATATATAGT 480
Db |||||
20420 GAAATTTATCTTAAGATACGTAAATGATTAATTTTGTACATTTGGAAATATATAGT 20479
QY |||||
481 TGTAGCCTGCGCGGTGCTCACGCTGTAAATCCAGCCCTTTGGAGGCCAGAGTGG 540
Db |||||
20480 TGTAGCCTGCGCGGTGCTCACGCTGTAAATCCAGCCCTTTGGAGGCCAGAGTGG 20539
QY |||||
541 CAGATCATAGAGTCAGAGATTTGAGACCAAGCTAGCCATATGTGAAAACCCCTCTCTA 600
Db |||||
20540 CAGATCATAGAGTCAGAGATTTGAGACCAAGCTAGCCATATGTGAAAACCCCTCTCTA 20599
QY |||||
601 CTAAAGATCAAAAAATTTAGCCAGAGTGTGAGCGACACTGTAAACCCAGCTGCTCA 660
Db |||||
20600 CTAAAGATCAAAAAATTTAGCCAGAGTGTGAGCGACACTGTAAACCCAGCTGCTCA 20659
QY |||||
661 GAGAGTGAGGCGAGAGAAATTTGTAACCCAGAGGCGAGAGGTGCGAGCGAAGCATCA 720
Db |||||
20660 GAGAGTGAGGCGAGAGAAATTTGTAACCCAGAGGCGAGAGGTGCGAGCGAAGCATCA 20719
QY |||||
721 CACCAATGCACTGAGCCCTGAGTGAAGAGGCGAAAGATCTCAACTCAAAAAAATTTT 780
Db |||||
20720 CACCAATGCACTGAGCCCTGAGTGAAGAGGCGAAAGATCTCAACTCAAAAAAATTTT 20779
QY |||||
781 AAGAAATATGTAAAGTTGTCTATACAAAATTAATAGCAGTGAAGCAAAAGTCTAAAG 840
Db |||||
20780 AAGAAATATGTAAAGTTGTCTATACAAAATTAATAGCAGTGAAGCAAAAGTCTAAAG 20839
QY |||||
841 CCTATGACCAATGTAACTAGAAATCTGTGGGAAACATTAATAGGAACTTAACCCAGT 900
Db |||||
20840 CCTATGACCAATGTAACTAGAAATCTGTGGGAAACATTAATAGGAACTTAACCCAGT 20899
QY |||||
901 CCTGGAAGTAAAGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAA 960
Db |||||
20900 CCTGGAAGTAAAGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAA 20959
QY |||||
961 AAAAAATTAATAATCCAGTTTACGTGTGAGAGAAATGGGATAGGGACTTAAGAGAAAT 1020
Db |||||
20960 AAAAAATTAATAATCCAGTTTACGTGTGAGAGAAATGGGATAGGGACTTAAGAGAAAT 21019
QY |||||
1021 CAAATAGGAATGTTTCATGATGATGTTAAGGACCCTGAGTAAAGGTTGAAGACCATTAATTA 1080
Db |||||
21020 CAAATAGGAATGTTTCATGATGATGTTAAGGACCCTGAGTAAAGGTTGAAGACCATTAATTA 21079
QY |||||
1081 TCTGCAACATCCGGGGAACTTTTATTTTATGAGTAAATGCTTGGCAATTTAAATAGAGAGCA 1140
Db |||||
21080 TCTGCAACATCCGGGGAACTTTTATTTTATGAGTAAATGCTTGGCAATTTAAATAGAGAGCA 21139
QY |||||
1141 GAGAAATGTAGACAGTTGAGATTGAGTCAGAGTTGAAATTTCTGCCAGACATGTGAAAAGAAAG 1200
Db |||||
21140 GAGAAATGTAGACAGTTGAGATTGAGTCAGAGTTGAAATTTCTGCCAGACATGTGAAAAGAAAG 21199
QY |||||
1201 AGACAGGTAGGCAAGAGAGTTGAAGAGATTATCAAGACAGAAATTAATGTGCTGCCAGT 1260
Db |||||
21200 AGACAGGTAGGCAAGAGAGTTGAAGAGATTATCAAGACAGAAATTAATGTGCTGCCAGT 21259
QY |||||
1261 GGCATCTAGCTTGAATCTAATCTGAGAGAAAGAAAGTAAAGCAAGCTTGTGTAAGT 1320
Db |||||
21260 GGCATCTAGCTTGAATCTAATCTGAGAGAAAGAAAGTAAAGCAAGCTTGTGTAAGT 21319
QY |||||
1321 TATGAAGAGAGTGAAGGCTTCAAGAGACCTACAGGTGTTGATTAATTAAGAAATGATTTG 1380
Db |||||

21320 TATGAAGAGAGTGAAGGCTTCAAGAGACCTACAGGTGTTGATTAATTAAGAAATGATTTG 21379
Db |||||
QY 1381 GAGAAAGAAATTAATCTGTGAGAGAGTGAATTTTCAAGGCTTTGAGTGACTCTCAATATCCAGA 1440
Db |||||
21380 GAGAAAGAAATTAATCTGTGAGAGAGTGAATTTTCAAGGCTTTGAGTGACTCTCAATATCCAGA 21439
QY |||||
1441 CACTGTGCTAAATGCTTCAAGAGACATGATCCCTGCCCTCAAGGGAGCTTACAGCCAAAAAC 1500
Db |||||
21440 CACTGTGCTAAATGCTTCAAGAGACATGATCCCTGCCCTCAAGGGAGCTTACAGCCAAAAAC 21499
QY |||||
1501 AAGAGATTAAGAAATTAACACCAATTAATTAAGAGACATTTGTGTAATATCAAGAAAG 1560
Db |||||
21500 AAGAGATTAAGAAATTAACACCAATTAATTAAGAGACATTTGTGTAATATCAAGAAAG 21559
QY |||||
1561 AAATTCGATCTAGTACTGTAATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGAA 1620
Db |||||
21560 AAATTCGATCTAGTACTGTAATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGAA 21619
QY |||||
1621 GTTTCAGATCCGCCCGGCGGTGCTCAAGCTGTAAATCCAGACATTTGGGAGGCCGA 1680
Db |||||
21620 GTTTCAGATCCGCCCGGCGGTGCTCAAGCTGTAAATCCAGACATTTGGGAGGCCGA 21679
QY |||||
1681 GCGGGGTGATCAAGAGTCAGAGATCAAGACATCTGTGTTAACACGGTGAACCCCG 1740
Db |||||
21680 GCGGGGTGATCAAGAGTCAGAGATCAAGACATCTGTGTTAACACGGTGAACCCCG 21739
QY |||||
1741 TCTCTACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db |||||
21740 TCTCTACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 21799
QY |||||
1801 ACTCAGAGAGCTGAGGCGAGAGAAATGCGGTGAACCCGGAGGTAGAGTTTGGGTAGCCG 1860
Db |||||
21800 ACTCAGAGAGCTGAGGCGAGAGAAATGCGGTGAACCCGGAGGTAGAGTTTGGGTAGCCG 21859
QY |||||
1861 AGATCGCGCACTGCGCTTCAGCTGCGGCGACAGATGAATCTCGCTCTCAAAAAAATTA 1920
Db |||||
21860 AGATCGCGCACTGCGCTTCAGCTGCGGCGACAGATGAATCTCGCTCTCAAAAAAATTA 21919
QY |||||
1921 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
Db |||||
21920 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 21979
QY |||||
1981 AGCATGTTACAGGAGAGAGAAATGTTTTCAGCAAGAGTACAGAGTGAAGGAAATAGAGAT 2040
Db |||||
21980 AGCATGTTACAGGAGAGAGAAATGTTTTCAGCAAGAGTACAGAGTGAAGGAAATAGAGAT 22039
QY |||||
2041 ATGTTCAAGAGAGAGACCCAGAGTCATGTTGTTAGGGTTAAGAGAAACACAGTGT 2100
Db |||||
22040 ATGTTCAAGAGAGAGACCCAGAGTCATGTTGTTAGGGTTAAGAGAAACACAGTGT 22099
QY |||||
2101 TTGCAATCTCCAGGTTCCATTAATGTCGTTAATGAATCAATAGTGTAGAGAACCTGCA 2160
Db |||||
22100 TTGCAATCTCCAGGTTCCATTAATGTCGTTAATGAATCAATAGTGTAGAGAACCTGCA 22159
QY |||||
2161 TTTTAAAAAATGAATTAATGATGAGAGAGAAATGAATTAATTAAGATGATTAATTAAT 2220
Db |||||
22160 TTTTAAAAAATGAATTAATGATGAGAGAGAAATGAATTAATTAAGATGATTAATTAAT 22219
QY |||||
2221 TTGAAGAGCAAGTATATTTTCTGCACTTTTGTCTCCAAATTTGAATCTGATATATTT 2280
Db |||||
22220 TTGAAGAGCAAGTATATTTTCTGCACTTTTGTCTCCAAATTTGAATCTGATATATTT 22279
QY |||||
2281 TTATGTATGAGATGTAATTAATCCAGATCAATTAATTTCTTACTAGTGAAGTGAAGCAAA 2340
Db |||||
22280 TTATGTATGAGATGTAATTAATCCAGATCAATTAATTTCTTACTAGTGAAGTGAAGCAAA 22339
QY |||||
2341 AATCTTAAAGCAGTGGCTGTACTTCTTATTTTTCAGAGAGAAATCAAGATC 2400
Db |||||
22340 AATCTTAAAGCAGTGGCTGTACTTCTTATTTTTCAGAGAGAAATCAAGATC 22399
QY |||||
2401 TGAAGAGCAAAACATTTTCTGAGGTTATAGAACCAAGTTATGCCATTTGCTAAAAGTGA 2460
Db |||||
22400 TGAAGAGCAAAACATTTTCTGAGGTTATAGAACCAAGTTATGCCATTTGCTAAAAGTGA 22459

QY 2461 TTCTTAGTAAATTTCTTCCCACTAGTGCATCTGCACTTCTAGTCTGTGGCCTGA 2520
DB 22460 TTCTTAGTAAATTTCTTCCCACTAGTGCATCTGCACTTCTAGTCTGTGGCCTGA 22519
QY 2521 AATPACGAATATATATAGTGAACAGATTCACAAGCTGGGGAAATATATTTGGGTAGTG 2580
DB 22520 AATPACGAATATATATAGTGAACAGATTCACAAGCTGGGGAAATATATTTGGGTAGTG 22579
QY 2581 GGTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGGCAAGGTATGTGGCTCACACCT 2640
DB 22580 GGTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGGCAAGGTATGTGGCTCACACCT 22639
QY 2641 ATATATCCGACACTTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAG 2700
DB 22640 ATATATCCGACACTTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAG 22699
QY 2701 CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTAAATTCAGCTGAGA 2760
DB 22700 CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTAAATTCAGCTGAGA 22759
QY 2761 TGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTAGGTGGTGAATCGCTTGACA 2820
DB 22760 TGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTAGGTGGTGAATCGCTTGACA 22819
QY 2821 CAGAGATTGAGGCTTAAGGTGAGCCATGATCACAACCTGCACTCCAGCTTGAGTGAAC 2880
DB 22820 CAGAGATTGAGGCTTAAGGTGAGCCATGATCACAACCTGCACTCCAGCTTGAGTGAAC 22879
QY 2881 AGGAAGACCCCTGCTCCATAAAAAAGAAATGTGATTTATCTCTAGACATGACAG 2940
DB 22880 AGGAAGACCCCTGCTCCATAAAAAAGAAATGTGATTTATCTCTAGACATGACAG 22939
QY 2941 CATTAAGTCAATTAAGTTGAGTTGAGAAATATATATGATCAGAAATTTATATCACT 3000
DB 22940 CATTAAGTCAATTAAGTTGAGTTGAGAAATATATATGATCAGAAATTTATATCACT 22999
QY 3001 GTGTGCTGTAGATATATGAAGAAATTAAGAGATGAGTCAAGGATTCCTTAAGTG 3060
DB 23000 GTGTGCTGTAGATATATGAAGAAATTAAGAGATGAGTCAAGGATTCCTTAAGTG 23059
QY 3061 TTGTGTTGTTGTTTGAAGACAGATCTCTTTTGTATCCAGGCTAGATGCAAGTG 3120
DB 23060 TTGTGTTGTTGTTTGAAGACAGATCTCTTTTGTATCCAGGCTAGATGCAAGTG 23119
QY 3121 CAGTCAATGCTCACCGACCTCAAACTCCAGCCTCAATTAATCTTCCAGCTCGGCT 3180
DB 23120 CAGTCAATGCTCACCGACCTCAAACTCCAGCCTCAATTAATCTTCCAGCTCGGCT 23179
QY 3181 CCCAAAGTCTGGAATTAACAGTGTGAGCCAAAGGTTTATGATGTGTCTGGCTAGT 3240
DB 23180 CCCAAAGTCTGGAATTAACAGTGTGAGCCAAAGGTTTATGATGTGTCTGGCTAGT 23239
QY 3241 GCTCTCAAACTTCAATGAGCAGACAAAGTACCGGGAACTGTACTCAACAGCTGGGT 3300
DB 23240 GCTCTCAAACTTCAATGAGCAGACAAAGTACCGGGAACTGTACTCAACAGCTGGGT 23299
QY 3301 TAAGCTGAGCCTCTGATTTCTAATCATGATCAAGTGAAGTGTGCTGTCAAG 3360
DB 23300 TAAGCTGAGCCTCTGATTTCTAATCATGATCAAGTGAAGTGTGCTGTCAAG 23359
QY 3361 GCCAAGCACTGAGCAGCAAGATCTAAGTGAAGTGAATTAATCAAGGTGATTAATGT 3420
DB 23360 GCCAAGCACTGAGCAGCAAGATCTAAGTGAAGTGAATTAATCAAGGTGATTAATGT 23419
QY 3421 AGTGAATTAATTAAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
DB 23420 AGTGAATTAATTAAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 23479
QY 3481 GTATATGAGCATGAGAGGAAAGATGGGGCCAAACATTAATCTGTTTGTGTGACT 3540
DB 23480 GTATATGAGCATGAGAGGAAAGATGGGGCCAAACATTAATCTGTTTGTGTGACT 23539

QY 3541 GAGAGGAGAAATTTAGCTCTATTAACAGAAATPAGAGAAAGTTGTTGGAGGAAA 3600
DB 23540 GAGAGGAGAAATTTAGCTCTATTAACAGAAATPAGAGAAAGTTGTTGGAGGAAA 23599
QY 3601 GAGAGTCTGTTTTCAGACGTGTGAGTCCAGGTGAGACAGATCTCCAAAGGAAAT 3660
DB 23600 GAGAGTCTGTTTTCAGACGTGTGAGTCCAGGTGAGACAGATCTCCAAAGGAAAT 23659
QY 3661 GAGAGTGAAGCACTTAAGAAATCTGTGCTCAGAGAGGAGCTGTGAGCTGT 3720
DB 23660 GAGAGTGAAGCACTTAAGAAATCTGTGCTCAGAGAGGAGCTGTGAGCTGT 23719
QY 3721 AGATCTGAGGTCATCAGACATPAGATTAGAAAGCAAGAGTGGCAACCAAGAGAC 3780
DB 23720 AGATCTGAGGTCATCAGACATPAGATTAGAAAGCAAGAGTGGCAACCAAGAGAC 23779
QY 3781 AATATACAAAGAGAGAGAGATGATGAGACTTTTGCCTTTTATGATGAGAGAGA 3840
DB 23780 AATATACAAAGAGAGAGAGATGATGAGACTTTTGCCTTTTATGATGAGAGAGA 23839
QY 3841 ACAGGAATGAAGAAATGAAGGAAAGCAGCTGTGAGAAATGTAAGCATCTGAAAAA 3900
DB 23840 ACAGGAATGAAGAAATGAAGGAAAGCAGCTGTGAGAAATGTAAGCATCTGAAAAA 23899
QY 3901 ATACACACTGTATGGAATCAAGGGAAGAAATTTCAAGAAAGAGGATATGTGACA 3960
DB 23900 ATACACACTGTATGGAATCAAGGGAAGAAATTTCAAGAAAGAGGATATGTGACA 23959
QY 3961 GTATTACAGCATTCAGGAATPACGCTAAAGTCAATCTGTGACTGATGACCTTGTG 4020
DB 23960 GTATTACAGCATTCAGGAATPACGCTAAAGTCAATCTGTGACTGATGACCTTGTG 24019
QY 4021 ATTTGTAGGGAACAATTAATTAAGAAATTTATGTGGTATATGAGGCAAAAG 4080
DB 24020 ATTTGTAGGGAACAATTAATTAAGAAATTTATGTGGTATATGAGGCAAAAG 24079
QY 4081 GAAGAGTTATCAAAATCAACAGGTGGAGTGAAGGATGAGTCTCCAAAGGTGAGGA 4140
DB 24080 GAAGAGTTATCAAAATCAACAGGTGGAGTGAAGGATGAGTCTCCAAAGGTGAGGA 24139
QY 4141 TCAGTGAATGTGGGAAAGGGGACAGAGCATTCATGCGCATCCAGGCAAGCCCTCA 4200
DB 24140 TCAGTGAATGTGGGAAAGGGGACAGAGCATTCATGCGCATCCAGGCAAGCCCTCA 24199
QY 4201 GAAGCTTCATGAGAGTTCAAGTATCAGAGAGTCTCTGATCCCTTAATCTTCTGGCTT 4260
DB 24200 GAAGCTTCATGAGAGTTCAAGTATCAGAGAGTCTCTGATCCCTTAATCTTCTGGCTT 24259
QY 4261 TGCATAGGCTTCAATGTGTAGGCAATGATTAATTAATTAATTAATTAATTAATTAAT 4320
DB 24260 TGCATAGGCTTCAATGTGTAGGCAATGATTAATTAATTAATTAATTAATTAATTAAT 24319
QY 4321 AACCTTCAACCCCTCTCCCTCCCTAATCATGCTGTGTCTTCCAGTACAGCAGTCCCTA 4380
DB 24320 AACCTTCAACCCCTCTCCCTCCCTAATCATGCTGTGTCTTCCAGTACAGCAGTCCCTA 24379
QY 4381 TCCATAAGCTAACCAATGTGTGCGCAGCTATCAGTCAACTCTACAAAAAGACATCACTTTG 4440
DB 24380 TCCATAAGCTAACCAATGTGTGCGCAGCTATCAGTCAACTCTACAAAAAGACATCACTTTG 24439
QY 4441 GAGATTTTAAGATTTTGAAGTGGCTGCAAGAAATTAAGTGAAGATCAATTAATTAAT 4500
DB 24440 GAGATTTTAAGATTTTGAAGTGGCTGCAAGAAATTAAGTGAAGATCAATTAATTAAT 24499
QY 4501 TTCACATATCAACAGTGTGCTAATTTTATATCAGGCGCATTAATTAATGTTTTAAACAAAG 4560
DB 24500 TTCACATATCAACAGTGTGCTAATTTTATATCAGGCGCATTAATTAATGTTTTAAACAAAG 24559
QY 4561 AGGTGATTAATTCAGATTTTCTTTTATTAACCTTAACATGATGACAGTGTGTAATG 4620
DB 24560 AGGTGATTAATTCAGATTTTCTTTTATTAACCTTAACATGATGACAGTGTGTAATG 24619
QY 4621 ATTTGGATGAGGCAATCTTTTTTTTTTGAAGATGTAATTCCTCCGACCTACTTCTC 4680

```
Db 24620 ATGGAGATGAGGGCAATACCTTTTGTGAAATGTTATATATCCCTGACCTTCTC 24679
Qy 4681 CTTGTTTCTTCACTCTCTCCCTACTACACAGAAAATTCTCTCTACTCAT 4740
Db 24680 CTTGTTTCTTCACTCTCTCCCTACTACACAGAAAATTCTCTCTACTCAT 24739
Qy 4741 TCCCTAAGTCGTCGTCGTTAAAGTTCCAGCTTGACGTCGATATCAAGAAC 4800
Db 24740 TCCCTAAGTCGTCGTCGTTAAAGTTCCAGCTTGACGTCGATATCAAGAAC 24799
Qy 4801 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860
Db 24800 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 24859
Qy 4861 GCCGGAACCTGACGGGAAAAACAAGCTCAGCGATATTTTCTGATCTGATATTA 4920
Db 24860 GCCGGAACCTGACGGGAAAAACAAGCTCAGCGATATTTTCTGATCTGATATTA 24919
Qy 4921 ATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
Db 24920 ATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24979
Qy 4981 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5040
Db 24980 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25039
Qy 5041 GCTCCAGACTGTGTGGGCTTAAAGTAAAGCAATCTGATGATGATGATGATGAT 5100
Db 25040 GCTCCAGACTGTGTGGGCTTAAAGTAAAGCAATCTGATGATGATGATGATGAT 25099
Qy 5101 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160
Db 25100 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25159
Qy 5161 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
Db 25160 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25219
Qy 5221 TTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5280
Db 25220 TTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 25279
Qy 5281 ACCGAGATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
Db 25280 ACCGAGATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25339
Qy 5341 ACCTCGCTTCTCGGTTCAAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Db 25340 ACCTCGCTTCTCGGTTCAAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25399
Qy 5401 CAGGACACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 5460
Db 25400 CAGGACACCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 25459
Qy 5461 ATGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520
Db 25460 ATGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25519
Qy 5521 AAAGTGTGGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5580
Db 25520 AAAGTGTGGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 25579
Qy 5581 ATGGGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640
Db 25580 ATGGGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25639
Qy 5641 GCCTTGAGCTCTCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 5700
Db 25640 GCCTTGAGCTCTCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 25699
Qy 5701 CAATATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 5760
```

```
Db 25700 CAATATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 25759
Qy 5761 AAAAATTAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5820
Db 25760 AAAAATTAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 25819
Qy 5821 ATGATGCAAAATGCTGCAAAAGATTAAGAAATGATGATGATGATGATGATGATGAT 5880
Db 25820 ATGATGCAAAATGCTGCAAAAGATTAAGAAATGATGATGATGATGATGATGATGAT 25879
Qy 5881 GTCATGGAAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5940
Db 25880 GTCATGGAAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25939
Qy 5941 TATAGGGGATGATGATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 6000
Db 25940 TATAGGGGATGATGATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 25999
Qy 6001 GTGTGTGACACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6060
Db 26000 GTGTGTGACACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26059
Qy 6061 GGTTGGCTAGAGCTGTTCTCATGACACATGATGATGATGATGATGATGATGATGAT 6120
Db 26060 GGTTGGCTAGAGCTGTTCTCATGACACATGATGATGATGATGATGATGATGATGAT 26119
Qy 6121 TGCCCATCTTCTCAGTGAATGACAGAGAGGATGATGATGATGATGATGATGATGAT 6180
Db 26120 TGCCCATCTTCTCAGTGAATGACAGAGAGGATGATGATGATGATGATGATGATGAT 26179
Qy 6181 GAAAGATGATGAAAGATGTGG 6201
Db 26180 GAAAGATGATGAAAGATGTGG 26200
```

RESULT 2
AX571875 5332 bp DNA linear PAT 29-NOV-2002
LOCUS AX571875
DEFINITION Sequence 3 from Patent WO02061060.
ACCESSION AX571875
VERSION AX571875.1 GI:26004000
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Yan, C., Ketchum, K., di Francesco, V. and Beasley, E.M.
Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
Patent: WO 02061060-A 3 08-AUG-2002;
JOURNAL PE Corporation (NY) (US)
FEATURES
source
1. 53332
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 6201; DB 6; Length 53332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACCATGGGCTCAGGCAATTTTATGAAGCCAAATATACGCTTAAATATGATGTGAGC 60
Db 20000 CACCATGGGCTCAGGCAATTTTATGAAGCCAAATATACGCTTAAATATGATGTGAGC 20059
Qy 61 CTAATAATCCAGAGAATCCCTTTGTAAAGTTTGTACAAATAATTAATATGATGAGAG 120
Db 20060 CTAATAATCCAGAGAATCCCTTTGTAAAGTTTGTACAAATAATTAATATGATGAGAG 20119
Qy 121 TTAATAGTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

Db 20120 TTAATAGTCTAATGGAATGCTGAACCAAGACCAATATCAGGCTTAGCAAAAATGCGACA 20179
Qy 181 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAA 240
Db 20180 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAA 20239
Qy 241 ATGTAAAGCCCTCAGCAATCGAAGGACAGGTTTACAGCAATTTGATCAAAAAGAAAA 300
Db 20240 ATGTAAAGCCCTCAGCAATCGAAGGACAGGTTTACAGCAATTTGATCAAAAAGAAAA 20299
Qy 301 CCACAGGCGCTTCCCTTCCCACTACTGATGTAAGAGCTTCTATTTTCCATAGTAGT 360
Db 20300 CCACAGGCGCTTCCCTTCCCACTACTGATGTAAGAGCTTCTATTTTCCATAGTAGT 20359
Qy 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAAAGTCTGAAAGAAAGCTCCCATTCAAAG 420
Db 20360 AAATTTTCTAGATACAGCTTGTAGAGCTCAAAAGTCTGAAAGAAAGCTCCCATTCAAAG 20419
Qy 421 GAAATTTATCTTAAGTACTGTAAATGATATCTAATTTTGTACATTTGGAATATATAGT 480
Db 20420 GAAATTTATCTTAAGTACTGTAAATGATATCTAATTTTGTACATTTGGAATATATAGT 20479
Qy 481 TGTAGCCCTGCGCGGCTCAAGCTGTAAATCCAGCCCTTTGGAGGCGCAGAGTGG 540
Db 20480 TGTAGCCCTGCGCGGCTCAAGCTGTAAATCCAGCCCTTTGGAGGCGCAGAGTGG 20539
Qy 541 CAGATCATAGAGTCAAGAGCTTTGAGACAGCCCTAGCCAACTGTGTAACCCCGTCTCTA 600
Db 20540 CAGATCATAGAGTCAAGAGCTTTGAGACAGCCCTAGCCAACTGTGTAACCCCGTCTCTA 20599
Qy 601 CTAAAGATACAAAATAATGACAGGTGTGTGGCGGCAACCTGTAAACCCCACTGTCTCA 660
Db 20600 CTAAAGATACAAAATAATGACAGGTGTGTGGCGGCAACCTGTAAACCCCACTGTCTCA 20659
Qy 661 GAGAGTGAAGGAGAGAAATTTGCTTGAACCCAGAGGCAAGGTGCGAGCGCAAAAGTCA 720
Db 20660 GAGAGTGAAGGAGAGAAATTTGCTTGAACCCAGAGGCAAGGTGCGAGCGCAAAAGTCA 20719
Qy 721 CACCAATGCACTGTAGCTGTAGTACAGGCGCAAGCTCCAACTCAAAAAAATAAAAA 780
Db 20720 CACCAATGCACTGTAGCTGTAGTACAGGCGCAAGCTCCAACTCAAAAAAATAAAAA 20779
Qy 781 AAGAAATATGTAAGTGTGTCTATACAAATAATAGGCACTGAGAAAGCAAGTGTAAAG 840
Db 20780 AAGAAATATGTAAGTGTGTCTATACAAATAATAGGCACTGAGAAAGCAAGTGTAAAG 20839
Qy 841 CCTATGACCATGTAATCTAGAAATACTGTGGGAACATATTAAGGGAACTTAACCCAGT 900
Db 20840 CCTATGACCATGTAATCTAGAAATACTGTGGGAACATATTAAGGGAACTTAACCCAGT 20899
Qy 901 CCTGGAAGTAAAGTTTGGAAAAGAAATGTTTGAAGACAAAGGGTTAAAGAGTGAATA 960
Db 20900 CCTGGAAGTAAAGTTTGGAAAAGAAATGTTTGAAGACAAAGGGTTAAAGAGTGAATA 20959
Qy 961 AAAAAATTAATAACCACTTTAGCTGTGTGAGAAATGGATAGGAGCTTAATAGAGAAAT 1020
Db 20960 AAAAAATTAATAACCACTTTAGCTGTGTGAGAAATGGATAGGAGCTTAATAGAGAAAT 21019
Qy 1021 CAAATAGGAATGTTTATGATGTTTAAGAACCTGTGTAAGGGTGAAGCACTTACATTA 1080
Db 21020 CAAATAGGAATGTTTATGATGTTTAAGAACCTGTGTAAGGGTGAAGCACTTACATTA 21079
Qy 1081 TCTGCAACCATCGCGGACTTTTATGATGTTTAAGTAAATGCTTGGCAATTTAAATAGAGACA 1140
Db 21080 TCTGCAACCATCGCGGACTTTTATGATGTTTAAGTAAATGCTTGGCAATTTAAATAGAGACA 21139
Qy 1141 GAGAAATGTAAGAGTGTGATGATGATGAGTGAAGTCTGCGCAACATGTGAAGAAAG 1200
Db 21140 GAGAAATGTAAGAGTGTGATGATGATGAGTGAAGTCTGCGCAACATGTGAAGAAAG 21199
Qy 1201 AGACAGTGAAGAGAGATTAAGAGATTAATCAAGACAGAAAGTTAAATGCTGCGCACT 1260
Db 21200 AGACAGTGAAGAGAGATTAAGAGATTAATCAAGACAGAAAGTTAAATGCTGCGCACT 21259
Qy 1261 GGCATCTAGTCTGAGTCTTAATCTGAGGGAAGAGATGAAGATTAAGAGCTTGTCTAGT 1320
Db 21260 GGCATCTAGTCTGAGTCTTAATCTGAGGGAAGAGATGAAGATTAAGAGCTTGTCTAGT 21319
Qy 1321 TATGAAGAGAGTGAAGGCTTCAAGGACCTACAGGCTTGAATTAATAGAAATGATGG 1380
Db 21320 TATGAAGAGAGTGAAGGCTTCAAGGACCTACAGGCTTGAATTAATAGAAATGATGG 21379
Qy 1381 GAGAAAGATTAATCTGAGAGAGTGAAGTTCAGGCTTGAATTAATAGAAATGATGG 1440
Db 21380 GAGAAAGATTAATCTGAGAGAGTGAAGTTCAGGCTTGAATTAATAGAAATGATGG 21439
Qy 1441 CACTGTCTTAAATGCTTCAAGACATGATCCCTGCCCTCAAGGACTTACAGCAAAAC 1500
Db 21440 CACTGTCTTAAATGCTTCAAGACATGATCCCTGCCCTCAAGGACTTACAGCAAAAC 21499
Qy 1501 AAGAGTAAAGAAATACACCAATCTAATATAGGCACTGTGTAAGATTAATCAAGAAAG 1560
Db 21500 AAGAGTAAAGAAATACACCAATCTAATATAGGCACTGTGTAAGATTAATCAAGAAAG 21559
Qy 1561 AAATAGATCTAGTACTGTAGATGTCACAGGCACTAAAGATATCTTCTAGTTTCAAGAA 1620
Db 21560 AAATAGATCTAGTACTGTAGATGTCACAGGCACTAAAGATATCTTCTAGTTTCAAGAA 21619
Qy 1621 GTTTAGATCGGCGCGGCGGCTGCTCAAGCTGTAAATCCACAGCACTTTGGAGGCGCA 1680
Db 21620 GTTTAGATCGGCGCGGCGGCTGCTCAAGCTGTAAATCCACAGCACTTTGGAGGCGCA 21679
Qy 1681 GCGGCGGTGATACAAAGGTCAAGAGATCAAGACATCTCGGTAAACAGGTAAACCCCG 1740
Db 21680 GCGGCGGTGATACAAAGGTCAAGAGATCAAGACATCTCGGTAAACAGGTAAACCCCG 21739
Qy 1741 TCTCTACAAAATAATTAATAATTAAGCAAGCGGTGTGGCGGCTGTAGTCCAGCT 1800
Db 21740 TCTCTACAAAATAATTAATAATTAAGCAAGCGGTGTGGCGGCTGTAGTCCAGCT 21799
Qy 1801 ACTCAGAGGCTGAGGCAAGAGATGCGGTGAACCCGGAGGTAGAGTTCGTGAGCCG 1860
Db 21800 ACTCAGAGGCTGAGGCAAGAGATGCGGTGAACCCGGAGGTAGAGTTCGTGAGCCG 21859
Qy 1861 AGATGCGGCACTGCGGCTCCAGCTGCGGGAAGAGTGAAGTGGGTCAAAAAAAT 1920
Db 21860 AGATGCGGCACTGCGGCTCCAGCTGCGGGAAGAGTGAAGTGGGTCAAAAAAAT 21919
Qy 1921 AAAAAAATTAATAATCTTCAAGATCTTAAACATGCACTTCAAGCTTAAATAGAG 1980
Db 21920 AAAAAAATTAATAATCTTCAAGATCTTAAACATGCACTTCAAGCTTAAATAGAG 21979
Qy 1981 AGCATGTTACAGGAGAGAAATGTTTTCAGCAAGGTACAGAGTGAAGGAAATAGAGAT 2040
Db 21980 AGCATGTTACAGGAGAGAAATGTTTTCAGCAAGGTACAGAGTGAAGGAAATAGAGAT 22039
Qy 2041 ATGTTCAAGGAAGAGAGCCCGAGATGCTTTTGTAGGCTTAAGAGAAACAGAGTGT 2100
Db 22040 ATGTTCAAGGAAGAGAGCCCGAGATGCTTTTGTAGGCTTAAGAGAAACAGAGTGT 22099
Qy 2101 TTGCAATCTCAGGCTTCAATTAAGTGTGATTAATCAATATGATGTTAGCAACTGCA 2160
Db 22100 TTGCAATCTCAGGCTTCAATTAAGTGTGATTAATCAATATGATGTTAGCAACTGCA 22159
Qy 2161 TTTTAAAGAAATGAATTAATGAGTGAAGAGATGAAGAAATATTAAGCATGATTAAT 2220
Db 22160 TTTTAAAGAAATGAATTAATGAGTGAAGAGATGAAGAAATATTAAGCATGATTAAT 22219
Qy 2221 TTGAAGAGCAAGTATTAATTTTCTGCACTTTTGTCTCAATGTGAATCTTAAATTT 2280
Db 22220 TTGAAGAGCAAGTATTAATTTTCTGCACTTTTGTCTCAATGTGAATCTTAAATTT 22279
Qy 2281 TTATGTATGATGATGAATACAGATACATATATTTCTTACTGTAGACTGCAAGTCAAAA 2340
Db 22280 TTATGTATGATGATGAATACAGATACATATATTTCTTACTGTAGACTGCAAGTCAAAA 22339

```
QY 2341 AATCTTTAAAGCATGCGCTGCTCTAATCTTATTTTGCAGAGGAGAAATCCAAATC 2400
| | | | |
Db 22340 AATCTTTAAAGCATGCGCTGCTCTAATCTTATTTTGCAGAGGAGAAATCCAAATC 22399
QY 2401 TGAAGGAGAAACATTTTGCCTGAGGCTTATGAAACCAAGCTTATGCCATGCTTAAAGTGA 2460
| | | | |
Db 22400 TGAAGGAGAAACATTTTGCCTGAGGCTTATGAAACCAAGCTTATGCCATGCTTAAAGTGA 22459
QY 2461 TTCTTAGTTAAATTTCTTCCCATAGTGCATCTGCACTTCTAGTCTGTGTGCTGA 2520
| | | | |
Db 22460 TTCTTAGTTAAATTTCTTCCCATAGTGCATCTGCACTTCTAGTCTGTGTGCTGA 22519
QY 2521 AATACGAATATATTTAGTGAACAGATACACAAGTCTGGGAAATATTTGGGTAGGTG 2580
| | | | |
Db 22520 AATACGAATATATTTAGTGAACAGATACACAAGTCTGGGAAATATTTGGGTAGGTG 22579
QY 2581 GCTGAAGGCTCATTTTCTAAGAAATGTGGAACCTTATGGGAGGGTATGGGTGCTCACACT 2640
| | | | |
Db 22580 GCTGAAGGCTCATTTTCTAAGAAATGTGGAACCTTATGGGAGGGTATGGGTGCTCACACT 22639
QY 2641 AATAATTCAGCATTTTGGGAGGCCAAGTCAAGAAAGATCGCTTGAAACCCAAAGTTCAAGA 2700
| | | | |
Db 22640 AATAATTCAGCATTTTGGGAGGCCAAGTCAAGAAAGATCGCTTGAAACCCAAAGTTCAAGA 22699
QY 2701 CTAGCATGGGCAACATAGCAAGACTCATCTCTACAAAAATTTAAATTCAGCTGAGCA 2760
| | | | |
Db 22700 CTAGCATGGGCAACATAGCAAGACTCATCTCTACAAAAATTTAAATTCAGCTGAGCA 22759
QY 2761 TGGTGCATACGCTGTAGTCCCACTGACCTGAGGAAAGTATGGTGTGATGCTTGACA 2820
| | | | |
Db 22760 TGGTGCATACGCTGTAGTCCCACTGACCTGAGGAAAGTATGGTGTGATGCTTGACA 22819
QY 2821 CAGAGCTTTGAGGCTAAGTGAAGCCATGATCAACAATCTGCACTCCAGCTTGAAGTACAG 2880
| | | | |
Db 22820 CAGAGCTTTGAGGCTAAGTGAAGCCATGATCAACAATCTGCACTCCAGCTTGAAGTACAG 22879
QY 2881 AGGAAGACCTGTGCTTAAAAAAGAAATGTGATTTTATCTTTAGACATGATCACT 2940
| | | | |
Db 22880 AGGAAGACCTGTGCTTAAAAAAGAAATGTGATTTTATCTTTAGACATGATCACT 22939
QY 2941 CATTAGCATTTAAGTTGAGTTGAGAGAAAAATATATGATCAGAGAAATTTTATCACT 3000
| | | | |
Db 22940 CATTAGCATTTAAGTTGAGTTGAGAGAAAAATATATGATCAGAGAAATTTTATCACT 22999
QY 3001 GTGCTGTAGATATATGAAAGAAATTAAGACTAGAGTCCAGGATTCACCTTAAGTG 3060
| | | | |
Db 23000 GTGCTGTAGATATATGAAAGAAATTAAGACTAGAGTCCAGGATTCACCTTAAGTG 23059
QY 3061 TTTGTTTGTGTTTGTGAGACAGAGTCTTTTGTGTTACCCAGGCTAGAGTGAATGTG 3120
| | | | |
Db 23060 TTTGTTTGTGTTTGTGAGACAGAGTCTTTTGTGTTACCCAGGCTAGAGTGAATGTG 23119
QY 3121 CAGTCAATGAGCTCACCGACGCTCAAACTCCAGCTCAAAATTAATCTTCCAGCTCGGCT 3180
| | | | |
Db 23120 CAGTCAATGAGCTCACCGACGCTCAAACTCCAGCTCAAAATTAATCTTCCAGCTCGGCT 23179
QY 3181 CCCAAAGTGTGAATTAACAGGTGTGAGCCAAAGGTTTATGATGTGTGCTGGCTAAGT 3240
| | | | |
Db 22180 CCCAAAGTGTGAATTAACAGGTGTGAGCCAAAGGTTTATGATGTGTGCTGGCTAAGT 23239
QY 3241 GCTCTCAAACTTCAAGTGAAGCAAGTGAACGGGAACTGTGACTCAACAAAGTCTGGGT 3300
| | | | |
Db 23240 GCTCTCAAACTTCAAGTGAAGCAAGTGAACGGGAACTGTGACTCAACAAAGTCTGGGT 23299
QY 3301 TAAAGCTGAGCTCTGCAATCTTAACATGAGTCAAGCTGATGACAGATGCTGTCAAGA 3360
| | | | |
Db 23300 TAAAGCTGAGCTCTGCAATCTTAACATGAGTCAAGCTGATGACAGATGCTGTCAAGA 23359
QY 3361 GCCAAGCACTGAGCAGCAGAGATCTAGTGAACAATTAAGTATCAAGTTGATATATAT 3420
| | | | |
Db 23360 GCCAAGCACTGAGCAGCAGAGATCTAGTGAACAATTAAGTATCAAGTTGATATATAT 24419
QY 3421 AGTGACAATAAGATGAGAGGAATGTGAATAATCAGTAAACAAGAAAGTTCACCTCTTG 3480
| | | | |

Db 23420 AGTGACAATAAGATGAGAGGAATGTGAATAATCAGTAAACAAGAAAGTTCACCTCTTG 23479
QY 3481 GTAATGTGAGCATGAGAGGAGAAAGATGAGGAGCCAAACATACTGTGTTGTGACT 3540
| | | | |
Db 23480 GTAATGTGAGCATGAGAGGAGAAAGATGAGGAGCCAAACATACTGTGTTGTGACT 23539
QY 3541 GACGAGAGAAATTTGAGCTCTATTATACGAAATATAGAGAAAGATGTTGTTGAGAGAA 3600
| | | | |
Db 23540 GACGAGAGAAATTTGAGCTCTATTATACGAAATATAGAGAAAGATGTTGTTGAGAGAA 23599
QY 3601 GAGGAGTCTGTTTACAGCTGTGAGGATCCAGGTGAGACAGATCTCCAAAGGGAAT 3660
| | | | |
Db 23600 GAGGAGTCTGTTTACAGCTGTGAGGATCCAGGTGAGACAGATCTCCAAAGGGAAT 23659
QY 3661 GAGCAGTAGGCAACCTTAAAGGAAATCTGTCTCAGAAAGGAGCTGTGAGCTGAGCTGT 3720
| | | | |
Db 23660 GAGCAGTAGGCAACCTTAAAGGAAATCTGTCTCAGAAAGGAGCTGTGAGCTGAGCTGT 23719
QY 3721 AGATCTGAGGCTCATCAGACATAGAGTTTATAGAGCAAGAGATGAGCAACCAAGAGC 3780
| | | | |
Db 23720 AGATCTGAGGCTCATCAGACATAGAGTTTATAGAGCAAGAGATGAGCAACCAAGAGC 23779
QY 3781 AATAACAAAGAGAGAGAGACTGATGATGAGACTTTTGCTTTAGATGAGAGAGGA 3840
| | | | |
Db 23780 AATAACAAAGAGAGAGAGACTGATGATGAGACTTTTGCTTTAGATGAGAGAGGA 23839
QY 3841 ACAGAGAAATGAAGGAATGAAGGAGAGAGCTTTGATGAGATGAGCATCTGAAAAAAA 3900
| | | | |
Db 23840 ACAGAGAAATGAAGGAATGAAGGAGAGAGAGCTTTGATGAGATGAGCATCTGAAAAAAA 23899
QY 3901 ATACACATGCTGATGAGAGTCAAGGAGAGAAATTTCAAGAGAGAGGATATGAGGACA 3960
| | | | |
Db 23900 ATACACATGCTGATGAGAGTCAAGGAGAGAAATTTCAAGAGAGAGGATATGAGGACA 23959
QY 3961 GTATTACAGCATCAGGAATACAGCTTAAAGTCAATCTTGTGACTGATGACTTGTG 4020
| | | | |
Db 23960 GTATTACAGCATCAGGAATACAGCTTAAAGTCAATCTTGTGACTGATGACTTGTG 24019
QY 4021 ATTTGTAGGAGCACACTAATTAATTAAGAAATTTATGTTGGGTATATGAGGACAAG 4080
| | | | |
Db 24020 ATTTGTAGGAGCACACTAATTAATTAAGAAATTTATGTTGGGTATATGAGGACAAG 24079
QY 4081 GAAAGGTTATTCGAATCAAGACAGTGGAGTATGAGTATCCAGGTTGAGAGGA 4140
| | | | |
Db 24080 GAAAGGTTATTCGAATCAAGACAGTGGAGTATGAGTATCCAGGTTGAGAGGA 24139
QY 4141 TCAGTGAATGTGGAAGGAGGACAGAGCATCCATGCCATCCAGGCAAGCCACTTCA 4200
| | | | |
Db 24140 TCAGTGAATGTGGAAGGAGGAGGACAGAGCATCCATGCCATCCAGGCAAGCCACTTCA 24139
QY 4201 GAAGCTTCATAGAGTTCAGCTATCCAGAGGCTCTGTACCTTAATTTTCTGGGTT 4260
| | | | |
Db 24200 GAAGCTTCATAGAGTTCAGCTATCCAGAGGCTCTGTACCTTAATTTTCTGGGTT 24259
QY 4261 TGCATAGGCTTATGTTGAGGAGATGATTTAATTAAGTATGAGGCACTGTGATCACTT 4320
| | | | |
Db 24260 TGCATAGGCTTATGTTGAGGAGATGATTTAATTAAGTATGAGGCACTGTGATCACTT 24319
QY 4321 AACCTTCAACCCCTCTCCCTCCCTAATCATATGCTGTGCTTTTCCAGTGAACCACTTCA 4380
| | | | |
Db 24320 AACCTTCAACCCCTCTCCCTCCCTAATCATATGCTGTGCTTTTCCAGTGAACCACTTCA 24379
QY 4381 TCCATAAGCTAACCAATGTGTGCAAGCTATCAGTCACTCTACAAAAGACATCACTTTG 4440
| | | | |
Db 24380 TCCATAAGCTAACCAATGTGTGCAAGCTATCAGTCACTCTACAAAAGACATCACTTTG 24439
QY 4441 GAGATTCATAAGATTTTGAAGTGGCTGTGAGAAATTAAGTGAAGATCAATATATAT 4500
| | | | |
Db 24440 GAGATTCATAAGATTTTGAAGTGGCTGTGAGAAATTAAGTGAAGATCAATATATAT 24499
QY 4501 TTCACAATATCAAGTGTGCTATTTTATATCAGGCGCATTTAAATGTTTTAAACAAG 4560
| | | | |
```

Db 24500 TTCACAAATACAGTGGTCTATTTTATATACAGGCCCATTAATAGTTTAAACAAAG 24559
 QY 4561 AGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACGTGACAGTGGTGAATAG 4620
 Db 24560 AGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACGTGACAGTGGTGAATAG 24619
 QY 4621 ATTGGGATGAGGGCAATCTTTTATTTTGAATGTTATATTCCTCGAACCTTACTTCTC 4680
 Db 24620 ATTGGGATGAGGGCAATCTTTTATTTTGAATGTTATATTCCTCGAACCTTACTTCTC 24679
 QY 4681 CTGTGTTTCTTCTACCTCTCTCCCTACTCACAGAAAATTCTCTCCCTACTCAT 4740
 Db 24680 CTGTGTTTCTTCTACCTCTCTCCCTACTCACAGAAAATTCTCTCCCTACTCAT 24739
 QY 4741 TCCCTGAATGCTGATGCTGTTAAGGTTCCAGCCTTGAACGTGAGGCTAATGAAACAC 4800
 Db 24740 TCCCTGAATGCTGATGCTGTTAAGGTTCCAGCCTTGAACGTGAGGCTAATGAAACAC 24739
 QY 4801 AGTGTACAGATGTAGATGATGTTGGGAGAAAGTGACATATGAGACCAATTAATTA 4860
 Db 24800 AGTGTACAGATGTAGATGATGTTGGGAGAAAGTGACATATGAGACCAATTAATTA 24859
 QY 4861 GCCGGAACGACGGGAAAAAAGAGTCAAGGATATTTTCTGGAATGCTGAATTA 4920
 Db 24860 GCCGGAACGACGGGAAAAAAGAGTCAAGGATATTTTCTGGAATGCTGAATTA 24919
 QY 4921 ATGATGATGTTGCTCACTCTGTATAGAGATGAGAAAGAAAATTTATTTTGGAGAG 4980
 Db 24920 ATGATGATGTTGCTCACTCTGTATAGAGATGAGAAAGAAAATTTATTTTGGAGAG 24979
 QY 4981 ATACCATGAATTTGTTTATAGACATGCTAAGTTTGAAGTATTTGGATGACAGGCA 5040
 Db 24980 ATACCATGAATTTGTTTATAGACATGCTAAGTTTGAAGTATTTGGATGACAGGCA 25039
 QY 5041 GCTCAGACTGTGTTGGGCTTAAGTAAAGGCAATCTGAGTTGAGATTAAGATTTTGA 5100
 Db 25040 GCTCAGACTGTGTTGGGCTTAAGTAAAGGCAATCTGAGTTGAGATTAAGATTTTGA 25099
 QY 5101 AATCATCAGAAATACGTTGTTCAATTAGACACTGTCAAGTGGTAAAGATTAAGGAGAC 5160
 Db 25100 AATCATCAGAAATACGTTGTTCAATTAGACACTGTCAAGTGGTAAAGATTAAGGAGAC 25159
 QY 5161 ATGTGTAAAGTATTAACAGAGATCAAAAGCGAAACCTTAAGATTAACAATATGTTATTA 5220
 Db 25160 ATGTGTAAAGTATTAACAGAGATCAAAAGCGAAACCTTAAGATTAACAATATGTTATTA 25219
 QY 5221 TTTATTAATTTATTAATGTTTATTTTATTTTATTTTATTTTATTTATTTATTTTATTT 5280
 Db 25220 TTTATTAATTTATTAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 25279
 QY 5281 ACGGAGTCTGCTGCTGCTGCCAGAGCTGAGTGCAGTGGCGCAAACTCAGCTCACTGCA 5340
 Db 25280 ACGGAGTCTGCTGCTGCTGCCAGAGCTGAGTGCAGTGGCGCAAACTCAGCTCACTGCA 25339
 QY 5341 ACCTCCGCTTCTGCTGCTCAAGGAGCTCTGCTCCTCAGCCTCTCAAGTATGCTGGACTA 5400
 Db 25340 ACCTCCGCTTCTGCTGCTCAAGGAGCTCTGCTCCTCAGCCTCTCAAGTATGCTGGACTA 25399
 QY 5401 CAGGACCCACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 5460
 Db 25400 CAGGACCCACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 25459
 QY 5461 ATGTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520
 Db 25460 ATGTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25519
 QY 5521 AAAAGTCTGGGATTAACAGGTATGAGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580
 Db 25520 AAAAGTCTGGGATTAACAGGTATGAGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25579
 QY 5581 ATGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640
 Db 25580 ATGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25639

QY 5641 GCCTTGCCCTCTCAAAAGTTCTGGGATTAACATGTGAGTCCCTGGCCCTGGCCCAATAT 5700
 Db 25640 GCCTTGCCCTCTCAAAAGTTCTGGGATTAACATGTGAGTCCCTGGCCCTGGCCCAATAT 25699
 QY 5701 CAATATATTAAGTTTATAGAGATGAAACCTATGAAAGAAAGAGAGAGAGAGAGAGAGAG 5760
 Db 25700 CAATATATTAAGTTTATAGAGATGAAACCTATGAAAGAAAGAGAGAGAGAGAGAGAG 25759
 QY 5761 AAAAATTAAGAGATTTGGAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5820
 Db 25760 AAAAATTAAGAGATTTGGAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25819
 QY 5821 ATGATGCCAATGCTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 5880
 Db 25820 ATGATGCCAATGCTGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 25879
 QY 5881 GTCATGGGAAACCATGTAATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 5940
 Db 25880 GTCATGGGAAACCATGTAATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 25939
 QY 5941 TATAGGGGATGAATTTTGTATTTCAAGAGACAAACAGTTCCATTAATGCGAAGATCTTA 6000
 Db 25940 TATAGGGGATGAATTTTGTATTTCAAGAGACAAACAGTTCCATTAATGCGAAGATCTTA 25999
 QY 6001 GTGTGTACACAGGAGATTAAGTGTCTGAATGTGATTTGGAAGAGAGAGATTAAGATTAAG 6060
 Db 26000 GTGTGTACACAGGAGATTAAGTGTCTGAATGTGATTTGGAAGAGAGAGATTAAGATTAAG 26059
 QY 6061 GGTGGCTAGAGCTGTTCTCATGACACTAATGTCATGAGACACAGCTGATCCAAAG 6120
 Db 26060 GGTGGCTAGAGCTGTTCTCATGACACTAATGTCATGAGACACAGCTGATCCAAAG 26119
 QY 6121 TGCCCACTCTTCACTGATGATGACAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTA 6180
 Db 26120 TGCCCACTCTTCACTGATGATGACAGAGAGAGATTAAGATTAAGATTAAGATTAAGATTA 26179
 QY 6181 GAAAGTAAATGGAAGATGTGG 6201
 Db 26180 GAAAGTAAATGGAAGATGTGG 26200

RESULT 3
 AC007242/c 112748 bp DNA linear PRI 30-SEP-2000
 LOCUS
 DEFINITION Homo sapiens BAC clone RP11-182H9 from 2, complete sequence.
 AC007242
 AC007242.3 GI:5931465
 VERSION
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Sulston, J.B. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE
 AUTHORS Kozlowicz, A., Stoneking, T., Hawking, M. and Hawrysko, C.
 TITLE The sequence of Homo sapiens BAC clone RP11-182H9
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (07-APR-1999) Genome Sequencing Center, Washington
 MO 63108, USA
 REFERENCE
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-SEP-1999) Genome Sequencing Center, Washington

REFERENCE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 112748)
AUTHORS Watson,R.
TITLE Direct Substitution
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 28, 1999 this sequence version replaced gi:5001521.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0182H09

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,B.,
Tateno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is Rpl1-53H10, 200 bp overlap, the
clone sequenced to the right is Rpl1-309N8. Actual start of this
clone is at base position 100126 of Rpl1-53H10; actual end is at
112748 of Rpl1-182H9.

The reads used to determine the sequence at the region of base
positions 31840 to 112748 consists of numerous polymorphic base
changes represented from both NH0182H09 and Rpl1-309N8. It is
unknown which base calls belong to which clone.

FEATURES

source

1. 112748

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RPl1-182H9"

/clone_1lb="RPCI-11"

1033. 11331

/rpt_family="Alu"

3574. 3737

/rpt_family="MIR"

3814. 4009

/rpt_family="MER1_type"

4048. 4358

/rpt_family="Alu"

repeat_region

6038. 6457
/rpt_family="MALR"
repeat_region
6477. 6680
/rpt_family="L2"
repeat_region
6681. 6978
/rpt_family="Alu"
repeat_region
7386. 7542
/rpt_family="Alu"
repeat_region
7543. 7567
/rpt_family="(TTTGG)n"
repeat_region
7568. 7702
/rpt_family="Alu"
repeat_region
8575. 8596
/rpt_family="AT_rich"
repeat_region
8599. 9529
/rpt_family="L1"
repeat_region
9685. 9875
/rpt_family="MIR"
repeat_region
10125. 10238
/rpt_family="MIR"
repeat_region
10357. 10458
/rpt_family="MIR"
repeat_region
10484. 10784
/rpt_family="MER1_type"
repeat_region
10895. 11020
/rpt_family="Alu"
repeat_region
11835. 12166
/rpt_family="MIR"
repeat_region
12082. 12118
/db_xref="GI:1234078"
repeat_region
12210. 12521
/rpt_family="(CA)n"
repeat_region
13068. 13394
/rpt_family="Alu"
repeat_region
13429. 13479
/rpt_family="L2"
repeat_region
13511. 13812
/rpt_family="Alu"
repeat_region
13967. 14092
/rpt_family="L2"
repeat_region
14111. 14139
/rpt_family="AT_rich"
repeat_region
14206. 14515
/rpt_family="AT_rich"
repeat_region
14516. 14559
/rpt_family="Alu"
repeat_region
14723. 14767
/rpt_family="Alu"
repeat_region
15360. 15575
/rpt_family="AT_rich"
repeat_region
15691. 15734
/rpt_family="MIR"
repeat_region
16277. 16577
/rpt_family="L2"
repeat_region
16605. 16649
/rpt_family="Alu"
repeat_region
16793. 16904
/rpt_family="L2"
repeat_region
16919. 17003
/rpt_family="MIR"
repeat_region
19655. 19795
/rpt_family="MIR"
repeat_region
20648. 20707
/rpt_family="GA-rich"
repeat_region
20881. 21187
/rpt_family="L2"
repeat_region
22002. 22304
/rpt_family="Alu"
repeat_region
23542. 23659
/rpt_family="Alu"
repeat_region
23944. 24249
/rpt_family="L2"
repeat_region
25966. 26170
/rpt_family="Alu"
repeat_region

```
repeat_region /rpt_family="L1"
26171..26461
rpt_family="Alu"
repeat_region 26476..26808
/rpt_family="MAlR"
repeat_region 26805..27119
/rpt_family="Alu"
repeat_region 27120..27165
/rpt_family="MAlR"
repeat_region 27186..27530
/rpt_family="L1"
repeat_region 27858..27971
/rpt_family="Alu"
repeat_region 28007..28310
/rpt_family="Alu"
repeat_region 28872..29076
/rpt_family="MIR"
repeat_region 29415..29758
/rpt_family="L1"
repeat_region 30018..30265
/rpt_family="Alu"
misc_feature 30648..30739
note="match to EST AA436054 (NID:92140968) zu01c12.r1"
repeat_region 31065..31159
/rpt_family="MER1_type"

Query Match 98.8%; Score 6129.2; DB 9; Length 112748;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 23; Indels 3; Gaps 3;

QY 1 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACGCTTAAATNGATGAGAC 60
DB 53580 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACGCTTAAATNGATGAGAC 53521
QY 61 CTAATAACCCAGAGAACTCCCTTGTAGATTGTGAATAAATAATATATAGTAGAG 120
DB 53520 CTAATAACCCAGAGAACTCCCTTGTAGATTGTGAATAAATAATATATAGTAGAG 53461
QY 121 TTAATAGTTCTAATGAAATGTGAACCAAGCCATATCAGCGCTAGCAAAATGGCAGA 180
DB 53460 TTAATAGTTCTAATGAAATGTGAACCAAGCCATATCAGCGCTAGCAAAATGGCAGA 53401
QY 181 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTACGCGCTATATAGTCTAAAGAAA 240
DB 53400 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTACGCGCTATATAGTCTAAAGAAA 53341
QY 241 ATGTGAAGCGCCTCAGCATCTGAAGGACAGTGTACAGCAATGTATCAAAAGAAAA 300
DB 53340 ATGTGAAGCGCCTCAGCATCTGAAGGACAGTGTACAGCAATGTATCAAAAGAAAA 53281
QY 301 CCACAGGCGCTTCCCTTCCCCCATCTTATGATGTAAGCAGTCTTCAATTTCCATAGTAGT 360
DB 53280 CCACAGGCGCTTCCCTTCCCCCATCTTATGATGTAAGCAGTCTTCAATTTCCATAGTAGT 53221
QY 361 AAATTTTCTAGATATACGCTTGTAGAGCTTAAAGAGTCTGAAAGAAAGCTCCCATTCAGAG 420
DB 53220 AAATTTTCTAGATATACGCTTGTAGAGCTTAAAGAGTCTGAAAGAAAGCTCCCATTCAGAG 53161
QY 421 GAAATTTATCTAAGTACTGTAAATGATAGTAATTTTGTACATTTGTAATGTAATATATAGT 480
DB 53160 GAAATTTATCTAAGTACTGTAAATGATAGTAATTTTGTACATTTGTAATGTAATATATAGT 53101
QY 481 TGTAGAGCTGAGCGGTGCTCAGCGCTGTAATCCAGAGCTTTGGAGGCGCAGAGTGGG 540
DB 53100 TGTAGAGCTGAGCGGTGCTCAGCGCTGTAATCCAGAGCTTTGGAGGCGCAGAGTGGG 53041
QY 541 CAGATCATAGAGTCAAGAGCTTTGAGACCAAGCTTACGCAATAGTGAAGCCCGTCTCTA 600
DB 53040 CAGATCATAGAGTCAAGAGCTTTGAGACCAAGCTTACGCAATAGTGAAGCCCGTCTCTA 52981
QY 601 CTAAGATATCAAAAAATTGACAGAGTGTGGCGGACACTGTAAACCCAGCTGTCTGA 660
DB 52980 CTAAGATATCAAAAAATTGACAGAGTGTGGCGGACACTGTAAACCCAGCTGTCTGA 52921
```

```
QY 661 GAGAGTGAAGGACAGAGAAATGCTTGAATCCACAGAGGACAGAGGTGACGCAAGCAAGATCA 720
DB 52920 GAGAGTGAAGGACAGAGAAATGCTTGAATCCACAGAGGACAGAGGTGACGCAAGCAAGATCA 52861
QY 721 CACCAATGCACTGTAGCGCTGATGACAGGGCAAGACTCCCAATCAAAAAAAAAAAAAA 780
DB 52860 CACCAATGCACTGTAGCGCTGATGACAGGGCAAGACTCCCAATCAAAAAAAAAAAAAA 52801
QY 781 AAGAAATATGAAATGTTGCTATTAACAAATTAATAGCAGTGAAGCAAGTGTAAAG 840
DB 52800 AAGAAATATGAAATGTTGCTATTAACAAATTAATAGCAGTGAAGCAAGTGTAAAG 52741
QY 841 CCTATGACCATGTGTAATCTAGAAATCTGTGGGACACATATTAAGGAACTTAACCACT 900
DB 52740 CCTATGACCATGTGTAATCTAGAAATCTGTGGGACACATATTAAGGAACTTAACCACT 52681
QY 901 CCTGGAAGTAAAGTTTGGAAAGAAATGTTTGAAGGCAAAAGGTTTAAAGAGTGAAGAA 960
DB 52680 CCTGGAAGTAAAGTTTGGAAAGAAATGTTTGAAGGCAAAAGGTTTAAAGAGTGAAGAA 52621
QY 961 AAAAAATTAATATACAGTTTACCTGTGTGAGAAATGGATAGGAGCTAATAGAGAAAT 1020
DB 52620 AAAAAATTAATATACAGTTTACCTGTGTGAGAAATGGATAGGAGCTAATAGAGAAAT 52561
QY 1021 CAAATAGCAATGTTTCATGTAATGTTTAAAGACCTGTGTAAGGCTGAAGCACTTACATTA 1080
DB 52560 CAAATAGCAATGTTTCATGTAATGTTTAAAGACCTGTGTAAGGCTGAAGCACTTACATTA 52501
QY 1081 TCTGACCATATGCGGAGACTTTTATTTATGTAATGCTTGGCAATTAATATAGAGAGCA 1140
DB 52500 TCTGACCATATGCGGAGACTTTTATTTATGTAATGCTTGGCAATTAATATAGAGAGCA 52441
QY 1141 GAGAAATGACAGATGTAATGATGACAGATGTAATGTAATGTAATGTAATGTAATGTAAT 1200
DB 52440 GAGAAATGACAGATGTAATGATGACAGATGTAATGTAATGTAATGTAATGTAATGTAAT 52381
QY 1201 AGACAGTGAAGCAAGAGTGAAGATTAATCAAGACAGATTAATGCTGTGCGCACT 1260
DB 52380 AGACAGTGAAGCAAGAGTGAAGATTAATCAAGACAGATTAATGCTGTGCGCACT 52321
QY 1261 GGCATCTAGTCTGAGCTTAATCTGAGGAAAGAACTGAAGAAATCAAGCTTCTGATAGT 1320
DB 52320 GGCATCTAGTCTGAGCTTAATCTGAGGAAAGAACTGAAGAAATCAAGCTTCTGATAGT 52261
QY 1321 TATGAAGAGATGGAAGGCTTCAAGACCTACAGGCTGTAATTAATAGAAATGATG 1380
DB 52260 TATGAAGAGATGGAAGGCTTCAAGACCTACAGGCTGTAATTAATAGAAATGATG 52201
QY 1381 GAGAAAGAAATTAATCTGAGAGAGTGAATTTCAAGCTTGAAGTACTTCACATACAGA 1440
DB 52200 GAGAAAGAAATTAATCTGAGAGAGTGAATTTCAAGCTTGAAGTACTTCACATACAGA 52141
QY 1441 CACTGTGCTAATGCTTCAAGACATGATCCTGCTTCAAGGACCTTACACCCAAAAAC 1500
DB 52140 CACTGTGCTAATGCTTCAAGACATGATCCTGCTTCAAGGACCTTACACCCAAAAAC 52081
QY 1501 AAGAGATTAAGAAATCAACACCAATCTATTATAGGACCTTGTAGAAATCAAGAAAG 1560
DB 52080 AAGAGATTAAGAAATCAACACCAATCTATTATAGGACCTTGTAGAAATCAAGAAAG 52021
QY 1561 AAATACGATCTAGTACTGATAGTGAACGGCATCAAGATATCTTCAAGTTTCAAGAA 1620
DB 52020 AAATACGATCTAGTACTGATAGTGAACGGCATCAAGATATCTTCAAGTTTCAAGAA 51961
QY 1621 GTTTTCAATCGGCGGCGGCTGCTCAAGCTTGAATCCAGACATTTGGAGAGCCGA 1680
DB 51960 GTTTTCAATCGGCGGCGGCTGCTCAAGCTTGAATCCAGACATTTGGAGAGCCGA 51901
QY 1681 GCGGCTGATCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 1740
DB 51900 GCGGCTGATCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 51841
```


QY	1741	TTCTCTAAGAAAAATTTTAAAGCAAGCGTGGTGGCGGGCGCTGTAGTCCAGCT	1800
Db	51840	TTCTTACTTAAAAATATAAAAAATTTAGCCAGGCGTGGTGGCGGGCGCTGTAGTCCAGCT	51781
QY	1801	ACTCAGGAGCGCTGAGCAGAGATGGCGTGAACCCGGGAGGTAGATTGGCTGAGCGC	1860
Db	51780	ACTCAGGAGCGCTGAGCAGAGATGGCGTGAACCCGGGAGGTAGATTGGCTGAGCGC	51721
QY	1861	AGATCGCGCACTGCGCTCAAGCCTGGGCGACAGAGTGAAGCTGGCTTCGAAAAA	1920
Db	51720	AGATCGCGCACTGCACTCCAGCCTGGGCGACAGAGTGAAGCTGGCTTC-AAAAA	51662
QY	1921	AAAAAAAAAAAAAGTTTCAGATCTTAAACACTGATTTTCAACAGCTAGAAATGAG	1980
Db	51661	AAAAAAAAAAAAAGTTTCAGATCTTAAACACTGATTTTCAACAGCTAGAAATGAG	51602
QY	1981	AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGAGTAGGAAATAGAGAT	2040
Db	51601	AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGAGTAGGAAATAGAGAT	51542
QY	2041	ATGTTCAAGAAAGAGACCCCAAGTCATGTTTGTAGGGTTTAAAGAAACAAGTGT	2100
Db	51541	ATGTTCAAGAAAGAGACCCCAAGTCATGTTTGTAGGGTTTAAAGAAACAAGTGT	51482
QY	2101	TTTGCAATCTCCAGGTTCCATTAGTGGCTTATGAATCAATATGTTAGTGTAGCACTGCA	2160
Db	51481	TTTGCAATCTCCAGGTTCCATTAGTGGCTTATGAATCAATATGTTAGTGTAGCACTGCA	51422
QY	2161	TTTTTAAAAAATGAATATATGATGAGAGAGAAATGAAAAATTTAGCATGCACTTCACT	2220
Db	51421	TTTTTAAAAAATGAATATATGATGAGAGAGAAATGAAAAATTTAGCATGCACTTCACT	51362
QY	2221	TTGAAAGAGCAAGTATTATTTTCTGCAACTTTTGTCTCAATTGTAACGTACTTATATTT	2280
Db	51361	TTGAAAGAGCAAGTATTATTTTCTGCAACTTTTGTCTCAATTGTAACGTACTTATATTT	51302
QY	2281	TTTATGATGATGATGATACCAAGTACATATATTTTCTACTGTAGACGTGCACTAAAA	2340
Db	51301	TTTATGATGATGATGATACCAAGTACATATATTTTCTACTGTAGACGTGCACTAAAA	51242
QY	2341	AATCTTTAAAGCACTGGCGCTGTCTAACTTCTATTTTGCAGAGGAGAAATCAAGATC	2400
Db	51241	AATCTTTAAAGCACTGGCGCTGTCTAACTTCTATTTTGCAGAGGAGAAATCAAGATC	51182
QY	2401	TGAGAGACAACAATTTTGGCTGAGTTATGAAACCAAGCTTATGCCATTGCTAAAAAGTGA	2460
Db	51181	TGAGAGACAACAATTTTGGCTGAGTTATGAAACCAAGCTTATGCCATTGCTAAAAAGTGA	51122
QY	2461	TTCTTACTTAAAAATTTCTTTGCCACTAGTGCATACGTCACTTACTTGTGTGGCTCTGA	2520
Db	51121	TTCTTACTTAAAAATTTCTTTGCCACTAGTGCATACGTCACTTACTTGTGTGGCTCTGA	51062
QY	2521	AATACAGAAATATATTAGTGAACAGACATACACAAGTCTGGGAGAAATATATGGGTAGGTG	2580
Db	51061	AATACAGAAATATATTAGTGAACAGACATACACAAGTCTGGGAGAAATATATGGGTAGGTG	51002
QY	2581	GCTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGGCAAGGTTATGTGGCTCACACCT	2640
Db	51001	GCTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGGCAAGGTTATGTGGCTCACACCT	50942
QY	2641	ATTAATTCAGACATTGTTGGAGGCCAAGTCAAGAAAGATCGCTTGAACCCAAAGTTCAAGA	2700
Db	50941	ATTAATTCAGACATTGTTGGAGGCCAAGTCAAGAAAGATCGCTTGAACCCAAAGTTCAAGA	50882
QY	2701	CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTTAAAAATCAGCTAGCA	2760
Db	50881	CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTTAAAAATCAGCTAGCA	50822
QY	2761	TGTTGGCATACGCGCTGTAGTCCCACTTACCTAGTGGAAAGTAGTGGGTGATCCCTTGA-C	2819
Db	50821	TGTTGGCATACGCGCTGTAGTCCCACTTACCTAGTGGAAAGTAGTGGGTGATCCCTTGA-C	50762
QY	2820	ACAGAGTTTGAAGGCTAAGGTAGACCAATGATCAACAACCTGCACTCCAGTTGATGAGCA	2879

D	b	49681	AATACACACTGTCAATGGAAGTCAAGGGAAGAGAATTTCAAGAAAGGGATGTGTGAC	49622
Q	y	3960	AGATTTCACAGCATCAGGAATACAGCTAAAAGTCACTACTGTGATGCAATTGACCTTGTG	4019
D	b	49621	AGATTTCACAGCATCAGGAATACAGCTAAAAGTCACTACTGTGATGCAATTGACCTTGTG	49565
Q	y	4020	GATTTCGAGGGACACCTAATAATAAGAAATTTATGTGGGTATATGAGGCAAAA	4079
D	b	49562	GATTTCGAGGGACACCTAATAATAAGAAATTTATGTGGGTATATGAGGCAAAA	49503
Q	y	4080	GGAAGAGGTTATCCAAATCAAAAGAGGTGGAGATAGGGAATGCTCCAAAGGTGAGGC	4139
D	b	49502	GGAAGAGGTTATCCAAATCAAAAGAGGTGGAGATAGGGAATGCTCCAAAGGTGAGGC	49443
Q	y	4140	ATCAGTGAATGTGGGAAGGGGCAACAGACATCCATGCGCCATCCAGGCAAGCCACTCC	4199
D	b	49442	ATCAGTGAATGTGGGAAGGGGCAACAGACATCCATGCGCCATCCAGGCAAGCCACTCC	49388
Q	y	4200	AGAAGCTCCATGAGAGTTCAAGATTCAGAAAGTCTCTGTACCTTAATCTTTCTGGGTT	4259
D	b	49382	AGAAGCTCCATGAGAGTTCAAGATTCAGAAAGTCTCTGTACCTTAATCTTTCTGGGTT	49322
Q	y	4260	TTGATATGGGCTCATCTGTGTAGGCAATGATTTATTAACATATGGCCACTGTGTATCACT	4319
D	b	49322	TTGATATGGGCTCATCTGTGTAGGCAATGATTTATTAACATATGGCCACTGTGTATCACT	49263
Q	y	4320	TAACTTCAACCCCTCTCCCTCCCTCAATCATGCTGTGGCTTTCCAGTGAACAGTCCCT	4379
D	b	49262	TAACTTCAACCCCTCTCCCTCCCTCAATCATGCTGTGGCTTTCCAGTGAACAGTCCCT	49202
Q	y	4380	ATCTGAAGCTAACCCCAATGTCTGTGCAGCTATCAGTCAACTCTCAAAAAACATCACTTT	4439
D	b	49202	ATCTGAAGCTAACCCCAATGTCTGTGCAGCTATCAGTCAACTCTCAAAAAACATCACTTT	49143
Q	y	4440	GGAGATTCCTAAGATTTTAGAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATA	4499
D	b	49142	GGAGATTCCTAAGATTTTAGAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATA	49083
Q	y	4500	TTTCACAAATATCACAGTCGTGCTATTTTATATCAGGGGCCATTAAATGGTTTTAAACAA	4559
D	b	49082	TTTCACAAATATCACAGTCGTGCTATTTTATATCAGGGGCCATTAAATGGTTTTAAACAA	49022
Q	y	4560	GAGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACTGATGACAGTGTGATATA	4619
D	b	49022	GAGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACTGATGACAGTGTGATATA	48965
Q	y	4620	GATTGGATGAGGGCAATCTTTTTTTTGAATGTATATTCCTCCGACCTTACTTTCT	4679
D	b	48962	GATTGGATGAGGGCAATCTTTTTTTTGAATGTATATTCCTCCGACCTTACTTTCT	48903
Q	y	4680	CCCTGTTTTCTTCACTCTCTCTCCCTCACTCAACAGAAACCTTCTCTCCCTTACTCA	4739
D	b	48902	CCCTGTTTTCTTCACTCTCTCTCCCTCACTCAACAGAAACCTTCTCTCCCTTACTCA	48843
Q	y	4740	TTCCCTGAATGTGTGTCTGTAAAGGTTCCACCTTGCAGTGAAGGCTAATCAGAACCA	4799
D	b	48842	TTCCCTGAATGTGTGTCTGTAAAGGTTCCACCTTGCAGTGAAGGCTAATCAGAACCA	48783
Q	y	4800	CAGTGTACAGATGTGATGATGTGTGGAGAAAGTGGACAGATATATGAGCCAAATTACTT	4859
D	b	48782	CAGTGTGTACAGATGTGATGATGTGTGGAGAAAGTGGACAGATATATGAGCCAAATTACTT	48723
Q	y	4860	AGCCGGAACTGACGGGAAAAACAAGATCAGCAATTTTTTTCTGTGATCTGAGTTTAA	4919
D	b	48722	AGCCGGAACTGACGGGAAAAACAAGATCAGCAATTTTTTTCTGTGATCTGAGTTTAA	48663
Q	y	4920	AATGATGATGTGTGCCATTCATCTGTGTATAGATCAAGAAAGAAATTTATTTTGGAGA	4979
D	b	48662	AATGATGATGTGTGCCATTCATCTGTGTATAGATCAAGAAAGAAATTTATTTTGGAGA	48603
Q	y	4980	GATACCAATGAATGTGTGTTTTAGACATGCTAAGTTTGAAGTATATGGAATGTACAGGCG	5039
D	b	48602	GATACCAATGAATGTGTGTTTTAGACATGCTAAGTTTGAAGTATATGGAATGTACAGGCG	48543

QY	5040	AGCTCCAGACTGTGTGGGCTTAAAGTAAAGCAATCTGAGTTGGAGATTAAGATTTTG	509
Db	48542	AGCTCCAGACTGTGTGGGCTTAAAGTAAAGCAATCTGAGTTGGAGATTAAGATTTTG	48483
QY	5100	AAATCATCAGAAATCCGGTGTTCATTAGACACTGTCAAGGGTAAGATAGCTAAAGGAG	5159
Db	48482	AAATCATCAGAAATCCGGTGTTCATTAGACACTGTCAAGGGTAAGATAGCTAAAGGAG	48422
QY	5160	CATGTGTAGATGATTAACAGAAATCAAGACGAAACCTTAAGAAATACATATGTATT	5219
Db	48422	CATGTGTAGATGATTAACAGAAATCAAGACGAAACCTTAAGAAATACATATGTATT	48365
QY	5220	ATTATTAATTTATTAATGTTTATTTTAAATTTATTTTATTTATTTATTTTAA	5279
Db	48362	ATTATTAATTTATTAATGTTTATTTTAAATTTATTTTATTTTATTTATTTTAA	48303
QY	5280	GACGGGAGTCTCGCTCGCTGCCACGGCTGGAATGCAAGTGGCCAACTCAGTCACTGC	5339
Db	48302	GATGGGAGTCTCGCTCGCTGCCACGGCTGGAATGCAAGTGGCCAACTCAGTCACTGC	48243
QY	5340	AACTCCGCTTCCCGGGTTCAAGAGGACCTCTGCCTCAAGCTCTCAAGTAGCTGGAGT	5399
Db	48242	AACTCCGCTTCCCGGGTTCAAGAGGACCTCTGCCTCAAGCTCTCAAGTAGCTGGAGT	48183
QY	5400	ACAGGCAACCAACCACTCACTGACTAATTTTGTATTTTATAGAGACGGGGTTTCA	5459
Db	48182	ACAGGCAACCAACCACTCACTGACTAATTTTGTATTTTATAGAGACGGGGTTTCA	48122
QY	5460	CATGTGGCCAGGCTGGTCTTGAACCTTGAACCTTGAAGATTCACTGCTTGGCTTC	5519
Db	48122	CATGTGGCCAGGCTGGTCTTGAACCTTGAACCTTGAAGATTCACTGCTTGGCTTC	48063
QY	5520	CAAAGTCTGAGGATTAAGGATTAAGACACTGCGCCGACCTATTTTGTTTTATTA	5579
Db	48062	CAAAGTCTGAGGATTAAGGATTAAGACACTGCGCCGACCTATTTTGTTTTATTA	48002
QY	5580	GATGGGGTCTTGTCTATGTTGCCAGGCTGGTCTGAACTCTGGACTCAAGCAATCTCC	5639
Db	48002	GATGGGGTCTTGTCTATGTTGCCAGGCTGGTCTGAACTCTGGACTCAAGCAATCTCC	47942
QY	5640	TGCTTGGCTCTCAAAAGTTCTGGGATTACACATGTGAGTCCCTGGCTGGCAGAAAT	5699
Db	47942	TGCTTGGCTCTCAAAAGTTCTGGGATTACACGTGTGAGTCCCTGGCTGGCAGAAAT	47883
QY	5700	TCATATATTAATTAATTTAGTAAAGTAAAGCCTATGAAAGAAACACCCAGAGGGGAGAA	5759
Db	47882	TCATATATTAATTAATTTAGTAAAGTAAAGCCTATGAAAGAAACACCCAGAGGGGAGAA	47822
QY	5760	GAAAAATTTAGAGATTGTGAACCAAAAGAAAGAGTGCCTCAGAAAGAAAGGACTGTCT	5819
Db	47822	GAAAAATTTAGAGATTGTGTGAACCAAAAGAAAGAGTGCCTCAGAAAGAAAGGACTGTCT	47762
QY	5820	TATGATCCAAATGCTGCAAAAGATTAAGAAATTAAGAAATCCATTGGGTTTCAATAGAAA	5879
Db	47762	TATGATCCAAATGCTGCAAAAGATTAAGAAATTAAGAAATCCATTGGATTTCAATAGAAA	47703
QY	5880	AGTATATGGGAAACCATGTGTAAAAAAACCTTGTGATATGACACAACTGTCGAAAAAGATT	5939
Db	47702	AGTATATGGGAAACCATGTGTAAAAAAACCTTGTGATATGACACAACTGTCGAAAAAGATT	47643
QY	5940	TTATAGGGGGATGAAATTTTGTATTTCAAGAGGCAAAACGTTCCATATGCGCAAGATCT	5999
Db	47642	TTATAGGGGGATGAAATTTTGTATTTCAAGAGGCAAAACGTTCCATATGCGCAAGATCT	47582
QY	6000	AGTGTGTGCAACGAGAGTTAGTGTCTGAAGTGAATTGGAAGAGCAGATCATTTGAGCTG	6059
Db	47582	AGTGTGTGCAACGAGAGTTAGTGTCTGAAGTGAATTGGAAGAGCAGATCATTTGAGCTG	47522
QY	6060	AGTTTGGCTAGAGCTGTTCTCATATGACACTAATGTATGATGAGTCAACACTGTATCCAA	6119
Db	47522	AGTTTGGCTAGAGCTGTTCTCATATGAGGCACTAATGTATGATGAGTCAACACTGTATCCAA	47463

QY 6120 GTGCCACATCTTCATGAATGA CAGAGCGATTGAGATTCAGTGAACCGCTAA 6179
|||||
Db 47462 GTGCCACATCTTCATGAATGA CAGAGCGATTGAGATTCAGTGAACCGCTAA 47403
QY 6180 AGAAGATGAATGAGATGTGG 6201
|||||
Db 47402 AGAAGATGAATGAGATGTGG 47381
RESULT 4
AC026170 145212 bp DNA linear PRI 08-NOV-2002
LOCUS Homo sapiens chromosome 3 clone RP11-163D23 map 3p, complete
DEFINITION
AC026170 145212 bp DNA linear PRI 08-NOV-2002
LOCUS Homo sapiens chromosome 3 clone RP11-163D23 map 3p, complete
sequence.
AC026170
AC026170.5 GI:24796716
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,M., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,X., Wang,X., Wang,X.,
Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J., and Yang,H.
Unpublished
Chromosome 3p genomic sequence
TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (21-MAR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
3 (bases 1 to 145212)
Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,M., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,X., Wang,X., Wang,X.,
Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J., and Yang,H.
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (19-JUL-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
4 (bases 1 to 145212)
Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,M., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,X., Wang,X., Wang,X.,
Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J., and Yang,H.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (08-NOV-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On Nov 8, 2002 this sequence version replaced gi:14916135.
-----Genome Center
Center:Beijing Center
Website:htp://hg.19t.ac.cn
http://www.genomics.org.cn
Contact:hg@19t.ac.cn
-----Project Information
Center project name:1f project
Center clone name: RP11-163D23
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 5% of reads
Chemistry: Dye-terminator: Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 145864 bases at least Q40
Consensus quality: 147954 bases at least Q30
Consensus quality: 148116 bases at least Q20
Insert size: 145212; sum-of-ctrligs
Quality coverage: 5.82x in Q20 bases;sum-of-ctrligs

FEATURES
source
location/Qualifiers
1..145212
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-163D23"
ORIGIN
Query Match 6.8%; Score 419.8; DB 9; Length 145212;
Best Local Similarity 52.3%; Pred. No. 2.1e-84;
Matches 1329; Conservative 0; Mismatches 1152; Indels 58; Gaps 16;
QY 449 TACTAATTTTGTACATTTGGAATATTAATGTTTGTAGCTGGCGGTGCTACAGCT 508
|||||
Db 92524 TACTAATTTTGTACATTTGGAATATTAATGTTTGTAGCTGGCGGTGCTACAGCT 92583
QY 509 GTAATCCCAAGCCCTTTGGAGGCGCAGAGGCGAGATCATGAGTCAAGGATTGAGACC 568
|||||
Db 92584 GTAATCCCAAGCCCTTTGGAGGCGCAGAGGCGAGATCATGAGTCAAGGATTGAGACC 92643
QY 569 AGCTAGCCCAAGATGATGAAACCCCGTCTCTAATAAGATACAAATAATTAGCCAGTGT 628
|||||
Db 92644 ATCTGCGCTAACAAGTGAATCCATCTCTAATAATAATAATAATTAGCCAGTGT 92703
QY 629 GGTGGCGCACACCTGTATTAACCCAGCTGCTCGAGAGTGAAGCAGAGAAATTGTTGAAC 688
|||||
Db 92704 GGTGGCGCGTGTGTATGATCCCGCTACCTCGAAGGCTGAGCAGAGAAATGCGTGAAC 92763
QY 689 CAGAGAGGAGAG-GTGCGAGGAGCAAGATCAACCAATGACT---GTAGCTGGATG 744
|||||
Db 92764 CAGAGAGGAGAGCTTACAGTGAAGCAAGATGCGCAGCTGCACTCCGTCTGCGGCG 92823
QY 745 ACAGGCGAAGCTCCAACTCAAAAAAAAAAAAAAAAAAAAAAATATGTAATTGCTCTATA 804
|||||
Db 92824 ACAGAGCGAAACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAGACTAGAAAGTGTGGA 92883
QY 805 ACAATAATTAAGGAGTGAAGCAAGAAAGTCTTAAGCCTATATCAATGTGTAATAGAAAT 864
|||||
Db 92884 ATGTGATGTAGCATGAGGAGATTTTAAATGGAATATATAGCAGGACCGGTGCTCA 92943
QY 865 ACTGTGGGAACA-TAATTAAGGAACCAACCACTGCTGGAAGTAAGTTTGTGA--- 920
|||||
Db 92944 TGCCTGTATTCACAGCATTTTGGAGAGCTGAGGCGCAGATCACTGAGTTGGAGTT 93003
QY 921 AAGGAATGTTGAGACAAGGTTAAAGAGTGAAGAAATAATAATAATCAAGTTT 980
|||||

KEYWORDS HTGS PHASE1: HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 185321)

REFERENCE 1 Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,

REFERENCE 2 Federicp,i,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,

REFERENCE 3 Mao,J., Marathe,R., Morehouse,A.J., Oefner,P., Palm,C.J.,

REFERENCE 4 Ramirez,D., Wilhelm,Y., Yu,S. and Davis,R.W.

REFERENCE 5 Unpublished

REFERENCE 6 2 (bases 1 to 185321)

REFERENCE 7 Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,

REFERENCE 8 Federicp,i,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,

REFERENCE 9 Mao,J., Marathe,R., Morehouse,A.J., Oefner,P., Palm,C.J.,

REFERENCE 10 Ramirez,D., Wilhelm,Y., Yu,S. and Davis,R.W.

REFERENCE 11 Direct Submission

REFERENCE 12 Submitted (20-Mar-2000) DNA Sequencing and Technology Center,

REFERENCE 13 Stanford University, 855 California Avenue, Palo Alto, CA 94304,

REFERENCE 14 USA

REFERENCE 15 On Jul 7, 2000 this sequence version replaced gi:8927681.

REFERENCE 16 ----- Genome Center

REFERENCE 17 Center: Stanford DNA Sequencing and Technology Development

REFERENCE 18 Center

REFERENCE 19 Center code: SDSITDC

REFERENCE 20 Web site: <http://sequence-www.stanford.edu/group/human/>

REFERENCE 21 Contact: hum-info@sequence.stanford.edu

REFERENCE 22 ----- Project Information

REFERENCE 23 Center project name: 863

REFERENCE 24 Center Clone name: RP11-269P18

REFERENCE 25 ----- Summary Statistics

REFERENCE 26 Sequencing Vector: M13mp18; X02513

REFERENCE 27 Chemistry: Dye-Primer; 0% of reads

REFERENCE 28 Assembly: Dye-terminator Big Dye; 100% of reads

REFERENCE 29 Assembly program: Phrap; version 0.990319

REFERENCE 30 Consensus quality: 175261 bases at least Q40

REFERENCE 31 Consensus quality: 179538 bases at least Q40

REFERENCE 32 Consensus quality: 180820 bases at least Q20

REFERENCE 33 Insert size: 187094; agarose-fp

REFERENCE 34 Insert size: 183721; sum-of-contigs

REFERENCE 35 Quality coverage: 8.2x in Q20 bases; agarose-fp

REFERENCE 36 Quality coverage: 8.0x in Q20 bases; sum-of-contigs.

REFERENCE 37 * NOTE: This is a 'working draft' sequence. It currently

REFERENCE 38 * consists of 17 contigs. The true order of the pieces is

REFERENCE 39 * is not known and their order in this sequence record is

REFERENCE 40 * arbitrary. Gaps between the contigs are represented as

REFERENCE 41 * runs of N, but the exact sizes of the gaps are unknown.

REFERENCE 42 * This record will be updated with the finished sequence

REFERENCE 43 * as soon as it is available and the accession number will

REFERENCE 44 * be preserved.

REFERENCE 45 1 1657: contig of 1657 bp in length

REFERENCE 46 * 1 1658 1757: gap of unknown length

REFERENCE 47 * 1 1758 3546: contig of 1789 bp in length

REFERENCE 48 * 3547 3646: gap of unknown length

REFERENCE 49 * 3647 5277: contig of 1631 bp in length

REFERENCE 50 * 5278 5378: gap of unknown length

REFERENCE 51 * 5378 8396: contig of 3019 bp in length

REFERENCE 52 * 8397 8496: gap of unknown length

REFERENCE 53 * 8497 11351: contig of 2855 bp in length

REFERENCE 54 * 11352 11451: gap of unknown length

REFERENCE 55 * 11452 15441: contig of 3990 bp in length

REFERENCE 56 * 15442 15541: gap of unknown length

REFERENCE 57 * 15542 21194: contig of 5653 bp in length

REFERENCE 58 * 21195 21294: gap of unknown length

REFERENCE 59 * 21295 26148: contig of 4854 bp in length

REFERENCE 60 * 26149 33566: gap of unknown length

REFERENCE 61 * 33567 33666: contig of 7318 bp in length

REFERENCE 62 * 33667 41541: contig of 7875 bp in length

REFERENCE 63 * 41542 41641: gap of unknown length

REFERENCE 64 * 41642 49652: contig of 8011 bp in length

REFERENCE 65 * 49653 65062: gap of unknown length

REFERENCE 66 * 65062: contig of 15310 bp in length

[illegible]

```
misc_feature 1. .1657 /note="assembly_name:Contig31"
misc_feature 1758. .3546 /note="assembly_name:Contig35"
misc_feature 3647. .5277 /note="assembly_name:Contig36"
misc_feature 5378. .8396 /note="assembly_name:Contig37"
misc_feature 8497. .11351 /note="assembly_name:Contig38"
misc_feature 11452. .15441 /note="assembly_name:Contig39"
misc_feature 15542. .21194 /note="assembly_name:Contig40"
misc_feature 21295. .26148 /note="assembly_name:Contig41
clone end:SP6
vector_end:right"
misc_feature 26249. .33566 /note="assembly_name:Contig42"
misc_feature 33667. .41541 /note="assembly_name:Contig43"
misc_feature 41642. .49652 /note="assembly_name:Contig44"
misc_feature 49753. .65062 /note="assembly_name:Contig45
clone_end:r7"
misc_feature 65163. .80491 /note="assembly_name:Contig46"
misc_feature 80592. .101499 /note="assembly_name:Contig47"
misc_feature 101600. .129616 /note="assembly_name:Contig48"
misc_feature 129717. .152331 /note="assembly_name:Contig49"
misc_feature 152432. .185321 /note="assembly_name:Contig50"
```

ORIGIN

Query Match	6.8%	Score 419.8;	DB 2;	Length 185321;
Best Local Similarity	52.3%	Pred. NO. 2.1e-84;		
Matches 1329;	Conservative	0;	Mismatches 1152;	Indels 58; Gaps 16;

Accession	Sequence	Position
Oy	TACTAATTTTGTACATTTGGATATATATAAGTTGTAGCTGGCGCGGTGGCTCAACGCT	508
Oy	449 TACTAATTTTGTACATTTGGATATATATAAGTTGTAGCTGGCGCGGTGGCTCAACGCT	508
Db	59345 TACTAAAAATAAAAAATATAAAAATTTTATAAAAAAGCGGCGCGGTGGCTCAACGCT	59404
Oy	509 GTAATCCACAGCCCTTTGGAGGCGACAGTGGCGAGTCAATGAAGTCAGAGTTTGAAC	568
Oy	59405 GTAATCCACAGCCCTTTGGAGGCGACAGTGGCGAGTCAATGAAGTCAGAGTTTGAAC	568
Db	59405 GTAATCCACAGCCCTTTGGAGGCGACAGTGGCGAGTCAATGAAGTCAGAGTTTGAAC	59466
Oy	569 AGCTTACCCAAACATGGGAAACCCCGCTCTACTATAAGATACAAAATTTAGCCAGGTG	628
Oy	59465 AGCTTACCCAAACATGGGAAACCCCGCTCTACTATAAGATACAAAATTTAGCCAGGTG	628
Db	59465 ATCTGGCTAACACGGTGAAATCCATCTTACTATAAATTAACAAAATTTAGCCGGGTG	59524

Db 61682 TAAATAATACAAAAATTAGCTGGGATGTGGGCGCATGCTTGTATGCCAGCTACTCTGTA 61741
QY 2796 AGCTAGTGGGTGGA--TCGCTTGACACAGAGTTTGAAGCTTAAGGTGAGCCATGATCAC 2853
Db 61742 GGCTGAGGACAAAGAAATTCCTTGAACCGGGAGATTGAGAGTTGACAGTGAAGCCAAATGTT 61801
QY 2854 ACAACCTGACTCCAGCTTGAAGTACAGAGAGAAACCTCTCCCTTAAGAAAAAGAAAG 2913
Db 61802 GCCACCTGACCTCCAGCTGGGCGACAGAGAGAGACCTCTGTCTCAAAAAAAGAAAAA 61861
QY 2914 TGGATTTATTCCTTAC 2932
Db 61862 AAAAAGATGATCCCTTGGC 61880

RESULT 6
HSM808294 2001 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686G1080 (from clone DKFZp686G1080).
DEFINITION BX648147 GI:34367306
ACCESSION
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2001)
Ansoyge,W., Krieger,S., Regier,T., Rittmuller,C., Schwager,B., Mewes,H.W., Weill,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and Wiemann,S.
COMMENT The German Human CDNA Consortium
CONSTRM Direct Submision
TITLE Submitted (27-AUG-2003) MIPS, Ingolstaeder Landstr.1, D-85764 Neuherberg, GERMANY
JOURNAL
COMMENT
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .2001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686G1080"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: h1cc3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
1956. .1961
polyA_signal
polyA_site
1982

Query Match 6.7%; Score 412.8; DB 9; Length 2001;
Best Local Similarity 92.1%; Pred. No. 7.7e-83;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 60
Db 53 CACCAGGGGCTCGACATTTATGAAACCGAATATACAGCTTAAATGAAATGTGAC 112
QY 61 CTAATAATCCAGAAAGAACTCCCTTTGTAAGTTTGAACAAATAATTAATAGTAGAG 120
Db 113 CTAATAATCCAGAAAGCAACCCCTTTGTAAGTTTGAACAAATAATTAATAGTAGAG 172
QY 121 TTAATAGTTTAAATGAGTAGTGAACCCAGAGCCATATCAGCGCTAGCAAAATGCGAGA 180
Db 173 TAAATAGTTTAAATGAGTAGTGAACCCAGAGCCATATCAGCGCTAGCAAAATGCGAGA 232

QY 161 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAA 240
Db 223 ATTCAATATGACATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAA 292
QY 241 ATGTAAAGCGCCCTCAGGCATCTGAAGAGCAGTGTATCAGCAATGATCAAAAAAGAAAA 300
Db 239 ATATAAAGCTCCCTCAGCGCCCGGAGACAGTGTATCAGCAATGATCAAAAAAGAAAA 352
QY 301 CCACAGCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 360
Db 353 CCACAGCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 412
QY 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAGAGAGAGAGAGAGAGAG 420
Db 413 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAGAGAGAGAGAGAGAGAG 472
QY 421 GAAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAGAGAGAGAGAGAGAGAG 479
Db 473 GAAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAGAGAGAGAGAGAGAGAG 532
QY 480 TTGT 483
Db 533 TTGT 536

RESULT 7
HSM807657 2002 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686I21199 (from clone DKFZp686I21199).
DEFINITION BX647511 GI:34366668
ACCESSION
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2002)
Koehler,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and Wiemann,S.
COMMENT The German Human CDNA Consortium
CONSTRM Direct Submision
TITLE Submitted (27-AUG-2003) MIPS, Ingolstaeder Landstr.1, D-85764 Neuherberg, GERMANY
JOURNAL
COMMENT
FEATURES
source
1. .2002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686I21199"
/tissue_type="human fetal brain"
/clone_lib="686 (synonym: h1cc3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"
/dev_stage="fetal"
1948. .1975
polyA_signal
polyA_site
1975

Query Match 6.7%; Score 412.8; DB 9; Length 2002;
Best Local Similarity 92.1%; Pred. No. 7.7e-83;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 60


```

Db      47  CACCCAGGCGCTGGACATATTTATGAAACCGATATACGCTTAAATGAAATGGAGAC 106
Oy      61  CTAAATACCCAGAGAAAGTCCCTTGTGTAGATTTGTAAACAAATTAATATAGTAGAG 120
Db      107  CTAAATACCCAGAGAAAGTCCCTTGTGTAGATTTGTAAACAAATTAATATAGTAGAG 166
Oy      121  TTATATGTTCTTAATGGAATGTGTGAACCCAGAGCCATATCAGCGCTAGACAAATGGCAGA 180
Db      167  TAAATAGTTCTTAATGGAATGTGTGAACCCAGAGCCATATCAGCGCTAGACAAATGGCAGA 226
Oy      181  ATTCAATATATCATCAAGTATATCTTCAAGAGCTTACAGGCGCCNAAAGATGCTAAAGAAA 240
Db      227  ATTCATATAGCATCAAGTATATCTTCAAGAGCTTACAGGCGCCNAAAGATGCTAAAGAAA 286
Oy      241  ATGTGAAGAGCCCTCAGCATCTGAAGAGACAGTGTATACAGCAATTTGATCAAAAGAAAA 300
Db      287  ATATGAAATCTCCCTCAGCGCCGAGAGACAGTGTATACAGCAATTTGATCAAAAGAAAA 346
Oy      301  CCACAGGCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 360
Db      347  CCACAGGCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 406
Oy      361  AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
Db      407  AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 466
Oy      421  GAAATTTATCTTAAGATCTGTAAATGATGATCTAA-TTTTTGTATCATTTGGAATATATAG 479
Db      467  GAAATTTATCTTAAGATCTGTAAATGATGATCTAA-TTTTTGTATCATTTGGAATATATAG 526
Oy      480  TTGT 483
Db      527  TTGT 530

```

```

RESULT 8
BC008944      1846 bp      mRNA      linear      PRI 04-NOV-2003
LOCUS        Homo sapiens ubiquitin-conjugating enzyme E2 variant 1, mRNA (CDNA
DEFINITION   BC008944
ACCESSION   BC008944
VERSION      BC008944.2 GI:38014077
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1846)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.P., Caesvanc,T.L.,
Schetz,T.E., Brownstein,M.J., Urdin,T.B., Toshlyuki,S.,
Carinci,P., Prange,C., Raha,S.S., Loqueilano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallue,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE      22388257
PUBMED       12477932
REFERENCE     2 (bases 1 to 1846)

```

```

AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (29-May-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA.
REMARK       NIH-MGC Project URL: http://mgc.nhl.nih.gov
On Oct. 28, 2003 this sequence version replaced gi:14286289.
COMMENT      Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

FEATURES
source       1. 1846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2960953"
/tissue_type="muscle, rhabdomyosarcoma"
/clone_id="NIH_MGC_17"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
<1. 1846
/gene="UBE2V1"
/notes="synonym: UBEV1, CIR1, CROC-1, UEV1A, CROC1, UEV-1"
/db_xref="locusID:7335"
/db_xref="MIM:602995"
<1. 431
/gene="UBE2V1"
/codon_start=3
/product="UBE2V1 protein"
/protein_id="AAH08944.2"
/db_xref="GI:38014078"
/db_xref="locusID:7335"
/db_xref="MIM:602995"
/translacion="SGGVKVPRLRLLEBEQGVGDGVTSWGLDEDEMTLTTRMT
GMIIQPPRIYENRISLKEGPPKYPPEAPPPVPTVKINNMVNSNGVDPRAISV
LAKMNSYSIKVYLOELRLMMSKNNMLPPPEQGCYSN"
12. 428
/gene="UBE2V1"
/notes="OU con; Region: Ubiquitin-conjugating enzyme.
Protein destined for proteasome-mediated degradation may
be ubiquitinated. Ubiquitination follows conjugation of
ubiquitin to a conserved cysteine residue of UBC
homologues. TSG101 is one of several UBC homologues that
lacks this active site cysteine"
/db_xref="CDD:pfam00179"

ORIGIN
Query Match 5.5%; Score 402.6; DB 9; Length 1846;
Best Local Similarity 92.5%; Pred. No. 16e-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Oy      16  CAATTATGAAGCCAAATATACACCTTAAATGATGTGACCTAATAATCCAGAG 75
Db      160  CAATTATGAAGCCAAATATACACCTTAAATGATGTGACCTAATAATCCAGAG 219
Oy      76  AACTCCCTTTGTAAGATTGTGAACAAATTAATATGATGAGTAAATAGTTCTTAATG 135
Db      220  CACCCCTTTGTAAGATTGTGAACAAATTAATATGATGAGTAAATAGTTCTTAATG 279
Oy      136  GAATGTGAAGCCAGAGCATATCAGCGCTGCAAAATGCGAATTCATATATCATCA 195

```

Db 280 GAGTGTGAGACCCAGAGCCATATCACTGCTAGCAAAAATGCGAAGTTCATATATGACATCA 339

Qy 196 AAGTTATCCCTTCAAGAGCTTCAAGCGCTATATGCTCTAAAGAAATGTGAAACGCGCTC 255

Db 340 AAGTTGCTCTGCAAGAGCTTCCGCGCTATATGCTCTAAAGAAATATGAAATCTCCCTC 399

Qy 256 AGCCATCTGAGAGACAGTGTATACAGCATTTGATCAAAAAAGAAACCAACGCGCTTCCC 315

Db 400 AGCCGCGCCGAGAGACAGTGTATACAGCATTTAATCAAAAAAGAAACCAACGCGCTTCCC 459

Qy 316 CTTCCCGCCCACTGTTAGTATGAGAGCTTCATTTCCATATGATAGTAAATTTTCTAGATAC 375

Db 460 CTTCCCGCCCACTGTTAGTATGAGAGCTTCATTTCCATATGATAGTAAATTTTCTAGATAC 519

Qy 376 AGCTTATAGAGCTCAAAAGTACTGAGAAAGAAAGCTCCCATTCAGAAATTTATCTTAAG 435

Db 520 GTCTTGTAGAGCTCAAAAGTACCGGAAAGAGAGCTCCCATTCAGAAATTTATCTTAAG 579

Qy 436 ATACTGTAATGATCTACTA-TTTTGTGACTTTGGAAATATATAGTTGT 483

Db 580 ATACTGTAATGATCTACTA-TTTTGTGACTTTGGAAATATATAGTTGT 628

RESULT 9
BC000468 1870 bp mRNA linear PRI 12-NOV-2003
LOCUS Homo sapiens ubiquitin-conjugating enzyme E2 variant 1, mRNA (cDNA
DEFINITION clone MGC:8586 IMAGE:2961017), complete cds.
ACCESSION BC000468
VERSION BC000468.2 GI:38197161
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1870)
Strasberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Frange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1870)
Strasberg, R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
On Nov 6, 2003 this sequence version replaced gi:12653396.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S.,
Dieckhoff, R.L., Granite, S., Guan, X., Gupta, J., Hsieh, P.,
Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantipod, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 1 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10863894.

FEATURES
source
1..1870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:8586 IMAGE:2961017"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NM_001086389.1"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1..1870
/gene="UBE2V1"
/note="synonyms: UBEV1, CIR1, CROC-1, UBEVIA, CROC1, UBEV-1"
/db_xref="LOCUSID:7335"
/db_xref="MIM:602995"
11..454
/gene="UBE2V1"
/codon_start=1
/product="UBE2V1 protein"
/protein_id="AAH00468.1"
/db_xref="GI:12653397"
/db_xref="LOCUSID:7335"
/db_xref="MIM:602995"
/translation="MAATGSGVAVPRNFRLLLEBEGQKGVGDGTVMGLDEDDMT
LRTWGMIGPRTYENRYSLEKCPKPEAPPFRFYTKIMNGVNSNGVDP
RLSTLAKWQNSYSIKVYLQELRLRMKMKKLPQPEGCYSN"
35..451
/gene="UBE2V1"
/note="UQ con; Region: Ubiquitin-conjugating enzyme.
Protein destined for proteasome-mediated degradation may
be ubiquitinated. Ubiquitination follows conjugation of
ubiquitin to a conserved cysteine residue of UBC
homologues. TSG101 is one of several UBC homologues that
lack this active site cysteine"
/db_xref="CD:pfam00179"

ORIGIN
Query Match 5.5%; Score 402.6; DB 9; Length 1870;
Best Local Similarity 92.5%; Pred. No. 1.6e-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
Qy 16 CAATTATGAAGCCAAATATACAGCTTAATAATGATGTGAGACTTAATACCAAGAG 75
Db 183 CAATTATGAAGCCAAATATACAGCTTAATAATGATGTGAGACTTAATACCAAGAG 242
Qy 76 AACTCCCTTTGTAGATTGTGAACAAATTAATATGATGAGTAACTACTTAAAG 135
Db 243 CACCCCTTTGTAGATTGTGAACAAATTAATATGATGAGTAACTACTTAAAG 302
Qy 136 GATGTGAAGCCAAAGCATATACAGCTTGAACAAATGAGAGATTGATATATCATCA 195
Db 303 GAGTGTGAGCCCAAGCATATACAGCTTGAACAAATGAGAGATTGATATATCATCA 362
Qy 196 AAGTTATCCCTTCAAGAGCTTCAAGCGCTATATGCTCTAAAGAAATGTGAAACGCGCTC 255

Db	363	AAATTGTCCTGCAGAGACTTCGGGGCTTAATGATGTCCTTAAAGAAATATGAAATCTCCCTC	422
Qy	256	AGGCATCTGAAGGACAGTGTACACGCAATTGATCAAAAAAGAAACCCACAGGCCCTCTCC	315
Db	423	AGCGGCGCGAAGGACAGTGTTCACGCAATTATCAAAAAAGAAACCCACAGGCCCTCTCC	482
Qy	316	CTTCCCCCANACTGATGTAAAGCATCTTCAATTTTCCATAGTAGTAAATTTTTCATGATAC	375
Db	483	CTTCCCCCNAATTCGATTTAACTCAAGTCTTCAATTTTCCACAGTAGTAAATTTTTCATGATAC	542
Qy	376	AGCTTGTAGACTCAAGTACTGGAAGAAAGATCCCATTCAAAGAAATTTATCTTTAAG	435
Db	543	GTCCTGTAGACTCAAGTACCGGAAAGAAAGATCCCATTCAAAGAAATTTATCTTTAAG	602
Qy	436	ATACTGTAAATGATATCTAA--TTTTTGTCAATTTGGAATATATTAAGTTGT	483
Db	603	ATACGTAAATGATATCTAATTTTGTGTGCATTTGGAATATATTAAGTTGT	651

LOCUS	AR000150	1930 bp	DNA	linear	PAT 04-DEC-1998
DEFINITION	Sequence 1 from patent US 5736331.				
ACCESSION	AR000150				
VERSION	AR000150.1				
KEYWORDS	GI:3962681				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1930)				
TITLE	Lin, S.-I. and Rochofsky, M. Lynn.				
JOURNAL	Method for identifying nucleic acids encoding c-fos promoter				
FEATURES	activating proteins				
source	Patent: US 5736331-A 1 07-APR-1998;				
	Location/Qualifiers				
	1..1930				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	6.5%; Score 402.6; DB 6; Length 1930;				
Best Local Similarity	92.5%; Pred. No. 1.6e-80;				
Matches	434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;				
QY	16 CAATTATGAAAGCCAAATATACAGCCTTAAATAGAAATGTGACCTTAATACCGAAG	75			
Db	245 CAATTATGAAAAACGAATATACAGCCTTAAATAGAAATGTGACCTTAATACCGAAG	304			
QY	76 AACCTCCCTTGTGAAGATTGTAACTAACTTAATATAGTAGTAATAGTTCTTAATG	135			
Db	305 CACCCCTTGTGAAGATTGTAACTAACTTAATATAGTAGTAATAGTTCTTAATG	364			
QY	136 GAATGGTGAACCCGAAGCCATATCAGCGCTAGCAAAATGGCAAGATTCATATATCATCA	195			
Db	365 GAGTGGTGAACCCGAAGCCATATCAGCTAGCAAAATGGCAAGATTCATATATCATCA	424			
QY	196 AAGTTATCTTCAAGAGCTTCAGCGCCCTTAATGATGTCTAAAGAAATGTGAACGCCCTC	255			
Db	425 AAGTTATCTTCAAGAGAGCTTCAGCGCCCTTAATGATGTCTAAAGAAATGTGAACGCCCTC	484			
QY	256 AGCCATCTGAAGAAGAGTGTACAGCAATTGATCAAAAAAGAAAAACAAGGCCCTTCCC	315			
Db	485 AGCCATCTGAAGAAGAGTGTACAGCAATTGATCAAAAAAGAAAAACAAGGCCCTTCCC	544			
QY	316 CTTCCCTCATCTGATGTAGAGCAGTCTTCAATTTTCCATAGTAGTAATTTTCTAGATAC	375			
Db	545 CTTCCCTCATCTGATGTAGAGCAGTCTTCAATTTTCCATAGTAGTAATTTTCTAGATAC	604			
QY	376 AGCTTGTAGAGCTCAAGTACTGGAAGAAAGCTCCCATCAAGGAATTTATCTTAAG	435			
Db	605 GTCTTGTAGAGCTCAAGTACTGGAAGAAAGCTCCCATCAAGGAATTTATCTTAAG	664			

Oy	436	ATTACTGAAATGATACATAA-TTTTGTGACATTGGATTAATTAAGTCT	483
Dd	665	ATTACTGTAATGATCACTAATTTTTTGTCATTTGAATATATAAGTCT	713
RESULT 11			
LOCUS	HSUJ9360	1982 bp	mRNA linear PRI 03-OCT-1997
DEFINITION	Homo sapiens DNA-binding protein (CROC-1A)	mRNA, complete cds.	
ACCESSION	U9360		
VERSION	U9360.1	GI:1066079	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1982)		
AUTHORS	Rochitsky,M.L. and Lin,S.L.		
TITLE	CROC-1 encodes a protein which mediates transcriptional activation of the human FOS promoter		
JOURNAL	Gene 195 (2), 141-149 (1997)		
MEDLINE	97449289		
PUBMED	9305758		
REFERENCE	2 (bases 1 to 1982)		
AUTHORS	Lin,S.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-OCT-1995) Stanley L. Lin, Tumor Biology, Scherer-Plogh Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..1962		

Query Match	Best Local Similarity	Matches	Score	DB	Length	Score	DB	Length
6.5%	92.5%	0	402.6	9	1982			
Conservative	0	Mismatches	34	Indels	1	Gaps	1	
43%								

ORIGIN	polyA_signal
Query Match	6.5%; Score 402.6; DB 9; Length 1982;
Best Local Similarity	92.5%; Pred. No. 1.6e-80;
Matches	43%; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
DB	311 CAATTATGAAAGCCAAATATATACAGCCTTAAATAATGAAATGTGACCTTAATACCCAGAG 75
DB	311 CAATTATGAAAGCCAAATATATACAGCCTTAAATAATGAAATGTGACCTTAATACCCAGAG 370
DB	76 AACTCCCCCTTGTAGATTGTGTAACAAAAATTAATAGAGAGATTAAATGTTCTATG 135
DB	371 CACCCCCCTTGTAGATTGTGTAACAAAAATTAATAGATTGAGTTAAATGTTCTATG 430
DB	136 GAATGTGAACCCAAAGCCATATACAGCCTAGCAAAATGGCAGAAATTCATATATATCA 195
DB	431 GAGTGTGAACCCAAAGCCATATACAGCCTAGCAAAATGGCAGAAATTCATATATATCA 490
DB	196 AAGTTATCTTCAAGAGCTTCAGGCTTAATGATGTCTAAAGAAAATGTGAAGCCCTC 255
DB	491 AAGTTATCTTCAAGAGAGCTTCAGGCTTAATGATGTCTAAAGAAAATGTGAAGCCCTC 550

gene	CDS
gene	1. .1982
CDS	70. .582
	/gene="CROC-1A"
	/functions="transcriptional activation of c-fos
	proto-oncogene promoter"
	/product="DNA-binding protein"
	/codon_start=1
	/protein_id="AA872015.1"
	/db_xref="GI:1066080"
	/translation="MPREVASLYKSGSKSDSBSRLPRPHCKGSGSPGPFLLLEE
	EEQKGVGDGVTSWGLDEDDMTLTWTGMIIGEPRTIYENRIYSLIKIEGPKYPEAP
	PFVRFVTKIKNMGNVNSNGVDDPRAISVLAQMUNYSIKVYLDELRLRLMSKEMKLE
	QPEPGQCYSN"
	1959. .1964
	/gene="CROC-1A"

Qy	256	AGCGATCTGAAGGACACTGTTTACAGCAATTGATCAAAAAGAAAAACACAGGCCCTTCCC	315			
Db	551	AGCGCCCGGAAGGACAGGTGTTACGCAATTATCAAAAAGAAAAACACAGGCCCTTCCC	610			
Qy	316	CTTCCGCCATCTCTGATGTGAAGCAGCTCTTCATTTTCCATGTAGTAATTTTCTAGATAC	375			
Db	611	CTTCCGCCCATTTGATGATTTTATCATGCTCTTCATTTTCCAGAGTAGTAATTTTCTAGATAC	670			
Qy	376	AGCTTTGAGAGCTCAAAAGTACTGGAAAAAGAAAGCTCCCATTCAGAAAGAAATTTTCTTAAG	435			
Db	671	GCTCTTGTAAGACCTCAAAAGTACCGGAAAAAGAAAGCTCCCATTCAGAAAGAAATTTTCTTAAG	730			
Qy	436	ATTACTGTAATGATACTTAA-TTTTTGTACATTTTGAATATATATAAGTGT	483			
Db	731	ATTACTGTAATGATACTTAAATTTTGTTCATTTGCAATTTGAATATATATAAGTGT	779			
RESULT 12						
HSU39361						
LOCUS	HSU39361	2135 bp	mRNA linear PRI 03-OCT-1997			
DEFINITION	Homo sapiens DNA-binding protein (CROC-1B) mRNA, complete cds.					
ACCESSION	U39361					
VERSION	U39361.1	GI:1066081				
KEYWORDS	.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens (human)					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	1 (bases 1 to 2135)					
AUTHORS	Rochefeky,M.L. and Lin,S.L.					
TITLE	CROC-1 encodes a protein which mediates transcriptional activation of the human FOS promoter					
JOURNAL	Gene 195 (2), 141-149 (1997)					
MEDLINE	97449289					
PUBMED	9305758					
REFERENCE	2 (bases 1 to 2135)					
AUTHORS	Lin,S.L.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-OCT-1995) Stanley L Lin, Tumor Biology, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA					
FEATURES						
source	Location/Qualifiers					
	1..2135					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="CROC-1B"					
	/tissue_type="Brain"					
gene	1..2135					
	/gene="CROC-1B"					
CDS	70..735					
	/gene="CROC-1B"					
	/function="transcriptional activation of c-fos proto-oncogene promoter"					
	/codon_start=1					
	/product="DNA-binding protein"					
	/protein_id="BAB72016.1"					
	/db_xref="GI:1066082"					
	/translation="MAKFKRHSPEALQLYPWECFVCLIRGTFPTNQIHKMSHYFGLPRWVTLIQDWHVILPRKRRHHVSHPHETFCITGVKVPNNRFLBEBEQDKVNGDYSWGIQDEDDMTLIRRWGMIIGPRITVENRIVSLKIEGPKPEAPFVRFTVATKINNGVNSNSGVDPRAISVLAKWQNSYSIKVVLQELIRLMSKENKKLPQPEGGQCVSN"					
	YSN"					
	2112..2117					
	/gene="CROC-1B"					
polYA_signal						
ORIGIN						
Query Match						
Best Local Similarity 92.5%; Pred. No.1.6e-80;						
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;						
16 CAATTATTGAAAGCCAAATATACAGCCTTAAATATGATGTGACCTTAATATCCAGAG 75						

Db	Accession	Source	Location/Qualifiers	Score	DB	Length	Indels	Gaps
Db	464	CAATTATGAAAACCGAATATACAGCTTAAATATGATGTGGACCTTAAATACCGAAG		6.5%	402.6	3802	1	1
Qy	76	AACCTCCCTTTGTAGATTGTGTAACAAAATTAATATAGATAGATTAAATAGTTCTAATG		6.5%	402.6	3802	1	1
Db	524	CACCCCTTTGTGTAGATTGTGTAACAAAATTAATATAGATAGATTAAATAGTTCTAATG		6.5%	402.6	3802	1	1
Qy	136	GAATGTGTAACCCAGAGCCATATCAGCGCTTACGAAATGCGAATTCATATTCATCA		6.5%	402.6	3802	1	1
Db	584	GAGGTGTGACCCAGAGCCATATCAGCGCTTACGAAATGCGAATTCATATTCATCA		6.5%	402.6	3802	1	1
Qy	196	AAGTTATCCTTCAAGACCTTCAAGCGCTTAAATGATGTCTAAAGAAATGTGAAAAGCCCTC		6.5%	402.6	3802	1	1
Db	644	AAGTTATCCTTCAAGACCTTCAAGCGCTTAAATGATGTCTAAAGAAATGTGAAAAGCCCTC		6.5%	402.6	3802	1	1
Qy	256	AGCATCTGAAGAGCAGTGTTCACGCAATTGATCAAAAAGAAAACACAGAGCCCTTCCC		6.5%	402.6	3802	1	1
Db	704	AGCGCCCGAAGAGCAGTGTTCACGCAATTGATCAAAAAGAAAACACAGAGCCCTTCCC		6.5%	402.6	3802	1	1
Qy	316	CTTCCCTTCAAGTGTGTAAGAGCAGTGTTCATTTTCCATAGTAAATTTTCTAGATAC		6.5%	402.6	3802	1	1
Db	764	CTTCCCTTCAAGTGTGTAAGAGCAGTGTTCATTTTCCATAGTAAATTTTCTAGATAC		6.5%	402.6	3802	1	1
Qy	376	AGCTTGTAGACTCAAGTGTGTAAGAGCAGTGTTCATTTTCCATAGTAAATTTTCTAGATAC		6.5%	402.6	3802	1	1
Db	824	AGCTTGTAGACTCAAGTGTGTAAGAGCAGTGTTCATTTTCCATAGTAAATTTTCTAGATAC		6.5%	402.6	3802	1	1
Qy	436	ATACGTAAATGATATCTAA-TTTTTGTACATTTGGAATATATTAAGTTGT		6.5%	402.6	3802	1	1
Db	884	ATACGTAAATGATATCTAA-TTTTTGTACATTTGGAATATATTAAGTTGT		6.5%	402.6	3802	1	1
RESULT 13	CO413441	Sequence 20512 from Patent WO0170979.		6.5%	402.6	3802	1	1
LOCUS	CO413441	Sequence 20512 from Patent WO0170979.		6.5%	402.6	3802	1	1
DEFINITION	CO413441	Sequence 20512 from Patent WO0170979.		6.5%	402.6	3802	1	1
ACCESSION	CO413441	Sequence 20512 from Patent WO0170979.		6.5%	402.6	3802	1	1
VERSION	CO413441.1	GI:41321222		6.5%	402.6	3802	1	1
KEYWORDS				6.5%	402.6	3802	1	1
SOURCE				6.5%	402.6	3802	1	1
ORGANISM				6.5%	402.6	3802	1	1
REFERENCE				6.5%	402.6	3802	1	1
AUTHORS				6.5%	402.6	3802	1	1
TITLE				6.5%	402.6	3802	1	1
JOURNAL				6.5%	402.6	3802	1	1
FEATURES				6.5%	402.6	3802	1	1
ORIGIN				6.5%	402.6	3802	1	1

```

Db      618 AATGTTCTCGAGAGCTTGGCGCTTAATGCTTAAGAAATATGAATCTCCCTC 677
Qy      256 AGCCATCTGAGAGAGAGTGTATACAGCAATTCATAAAGAAAAACACAGCCCTCC 315
Db      678 AGCCGCGCGAGAGAGAGTGTATACAGCAATTCATAAAGAAAAACACAGCCCTCC 737
Qy      316 CTTCCCGCATCTGATGTAAGAGAGTCTTCAATTTCCATAGTGTAAATTTTCTAGATAC 375
Db      738 CTTCCCGCGCATCTGATGTAAGAGAGTCTTCAATTTCCATAGTGTAAATTTTCTAGATAC 797
Qy      376 AGCTTGAGAGCTCAAGTACTGAGAAAGAAAGCTCCCATTCAGAAAGAAATTTATCTTAAG 435
Db      798 GTTTGTGAGAGCTCAAGTACTGAGAAAGAGAGCTCCCATTCAGAAAGAAATTTATCTTAAG 857
Qy      436 ATACTGTAATGATACTAA-TTTTGTACATTTGGAATATATAGTTGT 483
Db      858 ATACTGTAATGATACTAA-TTTTGTCCATTTGGAATATATAGTTGT 906

```

```

RESULT 14
AC148906      229392 bp      DNA      linear      HTG 21-APR-2004
LOCUS      Ocolemur garnettii clone CH256-538H20, WORKING DRAFT SEQUENCE.
DEFINITION
AC148906
VERSION      AC148906.1 GI:46430802
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Ocolemur garnettii (small-eared galago)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euleleostomi;
Mammalia; Eutheria; Primates; Strepsirrhini; Galagonidae; Ocolemur.

```

```

REFERENCE
AUTHORS      1 (bases 1 to 229392)
Antoniellis,A., Ayele,K., Benjamin,B., Blakeley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,
Larson,S., Lee-Lin,S.-Q., Legaepi,R., Maduro,Q.L., Maduro,V.B.,
Mayrullis,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddy-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
Stantford,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 229392)
Green,E.D.
Direct Submission
Submitted (21-APR-2004) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: SPY
Center clone name: 538H20
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 229306 bases at least Q40
Consensus quality: 229348 bases at least Q30
Consensus quality: 229382 bases at least Q20
Insert size: 247000; agarose-fp
Insert size: 229392; sum-of-contigs
Quality coverage: 10.87x in Q20 bases; sum-of-contigs
Quality coverage: 11.70x in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces

```

FEATURES
source
    * is believed to be correct as given, however the sizes
    * of the gaps between them are based on estimates that have
    * provided by the submittor.
    * This sequence will be replaced
    * by the finished sequence as soon as it is available and
    * the accession number will be preserved.
    1 229392: contig of 229392 bp in length.
        Location/Qualifiers
            1..229392
                /organism="Ocolemur garnettii"
                /mol_type="genomic DNA"
                /db_xref="taxon:10611"
                /clone="CH256-538H20"
                /clone_11b="CH256"
                /note="BAC resource: http://bacpac.chori.org/"
                1..229392
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:right"

```

```

ORIGIN
Query Match      6.2%; Score 387.2; DB 2; Length 229392;
Best Local Similarity 91.5%; Pred. No. 5.4e-77;
Matches 443; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

```

```

Qy      1 CACCATGGGCTCAGGCAATTTATGAAGCCAAATATACAGCTTAATATGAATGTGAC 60
Db      226746 CACCATGGGCTCAGGCAATTTATGAAGCCAAATATACAGCTTAATATGAAGGTGAC 226687
Qy      61 CTAATATCCAGAGAAGACCCCTTTGTAAGTTTGTAAGAAATTAATATGAGTAGAG 120
Db      226686 CTAATATCCAGAGAAGACCCCTTTGTAAGTTTGTAAGAAATTAATATGAGTAGAG 226627
Qy      121 TTAATAGTTCTAATGTAAGTGTGAACCCAGAGCATATCAGCGTAGCAAAATGCGACA 180
Db      226626 TTAATAGTTCTAATGTAAGTGTGAACCCAGAGCATATCAGCGTAGCAAAATGCGACA 226567
Qy      181 ATTATATATATCATCAAGTTATCTTCAAGAGCTTACGCGCTTAATGATGCTTAAGAAA 240
Db      226566 ATTATATATATCATCAAGTTATCTTCAAGAGCTTACGCGCTTAATGATGCTTAAGAAA 226507
Qy      241 ATGTGAAGCGCCTCAGCATCTGAAGGACAGTGTATACGAATGATCAAAAGAAAA 300
Db      226506 ATGTGAAGCGCCTCAGCATCTGAAGGACAGTGTATACGAATGATCAAAAGAAAA 226448
Qy      301 CCACAGGCGCTTCCCTTCCCATATCTGATGAAGAGAGTCTTCAATTTCCATAGTAGT 360
Db      226447 CCACAGGCGCTTCCCTTCCCATATCTGATGAAGAGAGTCTTCAATTTCCATAGTAGT 226389
Qy      361 AAATTTTCTAGATACAGTTGTGAGAGCTCAAGTACGAAAGAAAGAGCTCCCATTCAAAG 420
Db      226388 AAATTTTCTAGATACAGTTGTGAGAGCTCAAGTACGAAAGAGAGCTCCCATTCAAAG 226329
Qy      421 GAAATTTTCTAAGATCTGTAATGATACGA-TTTTGTGACATTTGGAATATATAG 479
Db      226328 GAAATTTTCTAAGATCTGTAATGATACGA-TTTTGTGACATTTGGAATATATAG 226269
Qy      480 TTGT 483
Db      226268 TTGT 226265

```

```

RESULT 15
AC008736      191925 bp      DNA      linear      PRI 27-SEP-2000
LOCUS      Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.
DEFINITION
AC008736
VERSION      AC008736.6 GI:10312244
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euleleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```


DB 76371 GAACCTTAAGAGCTCTATTAAAAATTAGAACTACATGAAAAATTAAAAAGCTGCTGA 76430
QY 2135 ATCAATATGGTGGTTAGCACTGCAATTTTAAAAATGAATTAATGATGAGAAAGAA 2194
DB 76431 AAGAAATTAAGAAAGACATTAATTAATGAAAGACATCCCACTTCATTGACTGAAAGAT 76490
QY 2195 TAGAAAAATTAGCATTGCACTTACATTTTGAAGAGCAAGATTAATTTTCTGCACTTTTG 2254
DB 76491 TTAATATCTTTAAGATGTAATCTTAATCAAGCAATCTACACATTCATTAATCCCTA 76550
QY 2255 CTCCAATTTAATCTGACTTAATTTTATGTAATGATGAAATACCAGATACATATATA 2314
DB 76551 CCAAAATCCACATATGATTTTTCAGAAATAGAAAAGCAGATTTCTAAATTCACATGAA 76610
QY 2315 TTTCTTACTGT---AGACTGAGTCAAAAATCTTTAAAGCAGTGGCTGCTTAACCTT 2370
DB 76611 ATCCCAAGGAGCCCAAAATAGCCAAAACATATCAAAAAAGAAACAAAGCTGGAGAGCTCA 76670
QY 2371 CCTTATTTTGCAGAGAGAAATCCAGATCTGAGAGACAAACATTTTGCCTGAGGTTAT 2430
DB 76671 CACTTGATACAAAACCTTTACTACAAAGCTACAGTAATCAAAAACAGTATGTACTGAAAT 76730
QY 2431 AGAACCAAGCTTATGCCATGTGTAAAGTATTCTTAGTTAAATCTTTCCCACTAGTGC 2490
DB 76731 AAAGACAGCATATAGACCAACAGAGTAGAATAGAGATCCGAATAAACCTTGCTGCTC 76790
QY 2491 CATACTGCACTTCTAGTTCTGTTGGCTGAAATACGAAATATATTAGTGAACAGCATAC 2550
DB 76791 TA-----GTCAATGATTTTGTGCAAGGGTACCAAAACGATT-----CAGGG 76832
QY 2551 ACAAGTCTGGGGAATATATTGGGTAGTGGCTGAGAGCTCAATTTCTTACGAATGTGG 2610
DB 76831 AAAGGTTTGTATCAAAATAGACACTGGGAAAACCTGATATTCAATGCAAAAGAAATG-AA 76891
QY 2611 ACCTTAGGCAAGGGTATGCTGCTCACACCTATATTCAGACCTTGGAGGCGCAAGTCA 2670
DB 76892 CTTGAGAGGCCAACAATAATGGCTCACACCTGTAATCCAGCACTTGGAAAGGCCAAGCA 76951
QY 2671 AGAAGATGCTTGAACCCAGAGTTCAAGACTAGCATGGGCAACATAGCAAGACCTCATC 2730
DB 76952 GGCAGATCACTGAGGTCAGAGTTGACACAAGCTGGCCAAATGATGTAATCCCGTC 77011
QY 2731 TCTACAAAAAATTAAAAATCAGCTGAGCATGCTGCTAGTCCCACTTACC 2790
DB 77012 TCTACTAAAAA-TACAAAAATTAGCCGGGTGTGTGTACATGCCCTTAATCCAGCTACT 77070
QY 2791 TGGGAAGCTAGGTGGGTGATGCTTGACACA--GGAGTTTGAGGCTAAGGTAGGCCATG 2848
DB 77071 CAGGAGGCTACGCCAAGAGATCACTTGAACCTGGAGGTGAGGTTGCAGTGAGCGCA 77130
QY 2849 ATCACAACAATGCACTCCAGCTTGAAGTGAAGAGAGAAACCTGTCTCTTAAAAAGAAAG 2908
DB 77131 ATTGACACCACTGCACTCCGGCTGGGTAAAGAGTGAAGCTGTCTCAAAAAAGAAAAATA 77190
QY 2909 AAA 2911
DB 77191 AAA 77193

Search completed: December 27, 2004, 04:18:15
Job time : 16943.5 secs

This Page Blank (uspto)

|||||
Db 241 CCTGTGTCTATATAAATAATAGCTGTGTCTTGGCAAGGCTCTGAGCTAGTACT 300
QY 301 CGGAACACTGAGGTGGAGGATCATGAGCCGAGAGGCTGAGGCTGAGTGAACAGTGA 360
Db 301 CGGAACACTGAGGTGGAGGATCATGAGCCGAGAGGCTGAGGCTGAGTGAACAGTGA 360
QY 361 TCACCCAGCTGATTCAGGCTGGAAGACAGAGGAGACCTGTCTTCCAAAAA 420
Db 361 TCACCCAGCTGATTCAGGCTGGAAGACAGAGGAGACCTGTCTTCCAAAAA 420
QY 421 AAAAAAATGCAAGAAAGACATCATTAATCTGACCTGGAGATTAATTTATGATG 480
Db 421 AAAAAAATGCAAGAAAGACATCATTAATCTGACCTGGAGATTAATTTATGATG 480
QY 481 AAATTCACATCTTTTAGAGAAATTAGCATTTCTGATTAATGATATATATATAT 540
Db 481 AAATTCACATCTTTTAGAGAAATTAGCATTTCTGATTAATGATATATATAT 540
QY 541 ATTATAAATTCAAATGGAATTAATAATCTGAGAAACTAGCTTCTCACTCTCAGTTGT 600
Db 541 ATTATAAATTCAAATGGAATTAATAATCTGAGAAACTAGCTTCTCACTCTCAGTTGT 600
QY 601 CAGTCAAAATTTAAATGTCTTTGGCCGGGTGGGTGGCTCAGGCTGTATCCGACGAC 660
Db 601 CAGTCAAAATTTAAATGTCTTTGGCCGGGTGGGTGGCTCAGGCTGTATCCGACGAC 660
QY 661 TTGGGAGGCGGAGCGGAGTGCATCAAGATTAGAGATCGAGACATCTGGCTAACA 720
Db 661 TTGGGAGGCGGAGCGGAGTGCATCAAGATTAGAGATCGAGACATCTGGCTAACA 720
QY 721 CGGTGAACCTGTCTCTACTAATAAATACAAAAATTAAGCCGGGTGGGTGGCCAGCC 780
Db 721 CGGTGAACCTGTCTCTACTAATAAATACAAAAATTAAGCCGGGTGGGTGGCCAGCC 780
QY 781 TTGATGCCAGCTGCTCAGAGGCTGAGGAGATGATGTGTGAACCCGGAGGCGGAG 840
Db 781 TTGATGCCAGCTGCTCAGAGGCTGAGGAGATGATGTGTGAACCCGGAGGCGGAG 840
QY 841 CTTGAGTGAAGCGAGATTCGCGCACTGCACCTCAGGCTGGGCGACAGTGCAGACTCTG 900
Db 841 CTTGAGTGAAGCGAGATTCGCGCACTGCACCTCAGGCTGGGCGACAGTGCAGACTCTG 900
QY 901 TCTCAAAAAAAGTTGAATGATGTCTTTGAGCCAGATGCTTCTTCTTTC 960
Db 901 TCTCAAAAAAAGTTGAATGATGTCTTTGAGCCAGATGCTTCTTCTTTC 960
QY 961 TTCTTCTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAATGGCT 1020
Db 961 TTCTTCTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAATGGCT 1020
QY 1021 TAAGTCTCTTCTTGTGGCAATTTGAAATGAAAAATACATGCTCATATTAATTAC 1080
Db 1021 TAAGTCTCTTCTTGTGGCAATTTGAAATGAAAAATACATGCTCATATTAATTAC 1080
QY 1081 CTGAACTTTTAAAAACATCATGAGGTTCAAAATCAAAATCTCAAAATATGTGT 1140
Db 1081 CTGAACTTTTAAAAACATCATGAGGTTCAAAATCAAAATCTCAAAATATGTGT 1140
QY 1141 GATATAGACATACTTATTTTCCCTTAATATGATTTGTTATATATCCCTCAATTC 1200
Db 1141 GATATAGACATACTTATTTTCCCTTAATATGATTTGTTATATATCCCTCAATTC 1200
QY 1201 TGTCTCACTTAT 1260
Db 1201 TGTCTCACTTAT 1260
QY 1261 AGCTTGAATTTGAAGCTACAGCTGAAACCAATTAGCCGATCTCTGACATGT 1320
Db 1261 AGCTTGAATTTGAAGCTACAGCTGAAACCAATTAGCCGATCTCTGACATGT 1320
QY 1321 GTGCTGTTAATATAGTACTTATATATATATATATATATATATATATATATATAT 1380
Db 1321 GTGCTGTTAATATAGTACTTATATATATATATATATATATATATATATATATAT 1380

1321 GTGCTGTTAATATAGTACTTATATATATATATATATATATATATATATATATAT 1380
QY 1381 CAAGCAAGAAAGAAAGATTTGATTTCAAACTCTCTCTCTCATACCATTTGGCTA 1440
Db 1381 CAAGCAAGAAAGAAAGATTTGATTTCAAACTCTCTCTCTCATACCATTTGGCTA 1440
QY 1441 ATATATATATTTGATCACTTTAAGAACATATAGGCTCACACATATTTTGAATAA 1500
Db 1441 ATATATATATTTGATCACTTTAAGAACATATAGGCTCACACATATTTTGAATAA 1500
QY 1501 TGAATGAATGGCAACCTTCTAAGACTATGGAATCACTATTTGTTGAAGCAAGAT 1560
Db 1501 TGAATGAATGGCAACCTTCTAAGACTATGGAATCACTATTTGTTGAAGCAAGAT 1560
QY 1561 GCAATGATATTTTCACTTTTCTGTGTTTATGATTTCTGTGTTCTTTGACTACTTA 1620
Db 1561 GCAATGATATTTTCACTTTTCTGTGTTTATGATTTCTGTGTTCTTTGACTACTTA 1620
QY 1621 AAGTTAGCTAGGTACCAATTTGTTTAAAGTGAACCAAAATGCTTCAATATAA 1680
Db 1621 AAGTTAGCTAGGTACCAATTTGTTTAAAGTGAACCAAAATGCTTCAATATAA 1680
QY 1681 GGTAGGAGAAAAATACCTCTCAACATGTCACCTTTAGCACCAAGAAAACTAATACA 1740
Db 1681 GGTAGGAGAAAAATACCTCTCAACATGTCACCTTTAGCACCAAGAAAACTAATACA 1740
QY 1741 TATCAACATCAATGATATCATTAATATCATGATGATGATGAAATGCTCAATCCCTTA 1800
Db 1741 TATCAACATCAATGATATCATTAATATCATGATGATGATGAAATGCTCAATCCCTTA 1800
QY 1801 AAACATATGATATCAATATGATCACTATGTCGCGCAAGAACCAAACTTAACTGCCAAA 1860
Db 1801 AAACATATGATATCAATATGATCACTATGTCGCGCAAGAACCAAACTTAACTGCCAAA 1860
QY 1861 TTTTATTTCTAATTAACAGCTGCTGTTTCAATGTTGCACTGATGATGCAAGCAA 1920
Db 1861 TTTTATTTCTAATTAACAGCTGCTGTTTCAATGTTGCACTGATGATGCAAGCAA 1920
QY 1921 TCCCTGTCTGATGATGATTTTCTGCACTGATGATGATGATGATGATGATGATGAT 1980
Db 1921 TCCCTGTCTGATGATGATTTTCTGCACTGATGATGATGATGATGATGATGATGAT 1980
QY 1981 GCTCAGAGGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 GCTCAGAGGAGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 AAAGGAGAGAAACAAGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
Db 2041 AAAGGAGAGAAACAAGATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
QY 2101 CTTGAACCTTCAAGATTTAGGCTCAAGAGCTGTGCAAAAGCTGATCAAGCTGATGCA 2160
Db 2101 CTTGAACCTTCAAGATTTAGGCTCAAGAGCTGTGCAAAAGCTGATCAAGCTGATGCA 2160
QY 2161 GCTGCTACATTTGTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Db 2161 GCTGCTACATTTGTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
QY 2221 CCAAGAGGCTGCTTCAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2280
Db 2221 CCAAGAGGCTGCTTCAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2280
QY 2281 AACCAAGAGGCTGAGCACTCCCTTTTGTAGCGGATCTGATTTCTTGTGCGGTAGGTCT 2340
Db 2281 AACCAAGAGGCTGAGCACTCCCTTTTGTAGCGGATCTGATTTCTTGTGCGGTAGGTCT 2340
QY 2341 AAACCAATATAATGAAATTTCTATTAAGTCAAGAAAAATTAATGCTGTGATTAATA 2400
Db 2341 AAACCAATATAATGAAATTTCTATTAAGTCAAGAAAAATTAATGCTGTGATTAATA 2400
QY 2401 TTTGGGAAATTTCTTGTAAACCAAAAGGAGAAATTAATCTTGTGCTTGGCTGACGAA 2460
Db 2401 TTTGGGAAATTTCTTGTAAACCAAAAGGAGAAATTAATCTTGTGCTTGGCTGACGAA 2460

Oy	2464	CTCACTGCTGCTGAAGTCGAGAAAGTGTCTCTCAAAATCTCTAAGGCTCTAAATTAC	2550
Dp	2461	ACTCACTTGCTGAGAGTCGAGAAAGTGTCTCTCAAAATCTCTAAGGCTCTAAATTAC	2520
Oy	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGTATTAAGTTGGTATGCTATGATTTGAAACTT	2580
Dp	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGTATTAAGTTGGTATGCTATGATTTGAAACTT	2580
Oy	2581	TAGTATTAAGTTCATGATTAATTAGCAATGCCATAGATTAATCCCTACAGCAATAAATTA	2640
Dp	2581	TAGTATTAAGTTCATGATTAATTAGCAATGCCATAGATTAATCCCTACAGCAATAAATTA	2640
Oy	2641	AGTGGACATGAAAAAAAAGCCAGACCTTAAACGAAAAAAGTTGCAAAAATCATCATCAA	2700
Dp	2641	AGTGGACATGAAAAAAAAGCCAGACCTTAAACGAAAAAAGTTGCAAAAATCATCATCAA	2700
Oy	2701	AGAGATTTAGGTTAACTGAAATGTTTAAAGACACTTTTTAGTGAGAAAGAAATGTAGTA	2760
Dp	2701	AGAGATTTAGGTTAACTGAAATGTTTAAAGACACTTTTTAGTGAGAAAGAAATGTAGTA	2760
Oy	2761	TTTTCAGAGATTTGATTAACCATTAATGCTCTTTTCAAGGATCTTTCAAGAAAAGTGCCTTTTG	2820
Dp	2761	TTTTCAGAGATTTGATTAACCATTAATGCTCTTTTCAAGGATCTTTCAAGAAAAGTGCCTTTTG	2820
Oy	2821	GGGGTACAGAAAGCTTTAGAAAAACATTTGAAAGTGAATAAGGCAAAATAAGAAAAAT	2880
Dp	2821	GGGGTACAGAAAGCTTTAGAAAAACATTTGAAAGTGAATAAGGCAAAATAAGAAAAAT	2880
Oy	2881	GGTTTTCACGAGCACTGAATCTTTTACTTTGATAAATTTTATTTTCTGCTCTTTCTTTT	2940
Dp	2881	GGTTTTCACGAGCACTGAATCTTTTACTTTGATAAATTTTATTTTCTGCTCTTTCTTTT	2940
Oy	2941	CTCTAGCTTAAAGACCTTAAAGAAAGCATCAATGCCAATGATTTTCAACCCAGGAGGA	3000
Dp	2941	CTCTAGCTTAAAGACCTTAAAGAAAGCATCAATGCCAATGATTTTCAACCCAGGAGGA	3000
Oy	3001	CTTCAAGCTGCGCCGTCGCCAGAAAGTTCAAAGAGTAAAGAGCCAGGAGTAAACATGATTTGT	3060
Dp	3001	CTTCAAGCTGCGCCGTCGCCAGAAAGTTCAAAGAGTAAAGAGCCAGGAGTAAACATGATTTGT	3060
Oy	3061	TTTTCAGAGAAAGATCTTGAGGCAAGGTTTTCAATGGGTGATGAGACATGATTTGATC	3120
Dp	3061	TTTTCAGAGAAAGATCTTGAGGCAAGGTTTTCAATGGGTGATGAGACATGATTTGATC	3120
Oy	3121	AAGAAGATTTAAATGTGACTTGTCTAGAGAGGCGCGCTTGTCTTCCAGGGCAATTAATCT	3180
Dp	3181	GAGGAGGCTTCCCAAGCTGCTCTCTGCGCAATGCTGTCTAAATTTTCCCTGGGAAAAAAAGT	3240
Oy	3181	GAGGAGGCTTCCCAAGCTGCTCTCTGCGCAATGCTGTCTAAATTTTCCCTGGGAAAAAAAGT	3240
Dp	3241	CAACACTTAAAAAAAAGTGTCTTTCTCTCTTCCCTTCAACCGCTCTTTTCCCAATTC	3300
Oy	3241	CAACACTTAAAAAAAAGTGTCTTTCTCTCTTCCCTTCAACCGCTCTTTTCCCAATTC	3300
Dp	3301	CAACACTTAAAAAAAAGTGTCTTTCTCTCTTCCCTTCAACCGCTCTTTTCCCAATTC	3300
Oy	3301	CCTAGAGAGAGAAAGAGCTTCCCTTTTGGGGGAGCCCTCACTTAATTTGAACCTTGGAGAA	3360
Dp	3361	CCTAGAGAGAGAAAGAGCTTCCCTTTTGGGGGAGCCCTCACTTAATTTGAACCTTGGAGAA	3360
Oy	3361	GCTGGGTGAAGGCTCTTAATGCGACAGTTTACAAAGGGATTTAGCAGGTGAGTGAACATAG	3420
Dp	3421	GCTGGGTGAAGGCTCTTAATGCGACAGTTTACAAAGGGATTTAGCAGGTGAGTGAACATAG	3420
Oy	3421	CTGGGAGAGACTTTAGAGATGAGTCCCGCCCCCAATTTCAATTAATAAGCCAGGT	3480
Dp	3421	CTGGGAGAGACTTTAGAGATGAGTCCCGCCCCCAATTTCAATTAATAAGCCAGGT	3480
Oy	3481	GAGACATCATTAAGAAAGTTTATAGACACTGAGACCTGTGCAACACCAATGGCCGACAGGGA	3540
Dp	3481	GAGACATCATTAAGAAAGTTTATAGACACTGAGACCTGTGCAACACCAATGGCCGACAGGGA	3540

OY	3541	GAGGACATGATTA	CTTAAACAGCCTTGAAGAAA	AAAAAACAACCGCCCTGCCTTAATTA	3600
Ds	3541	GAGGACATGATTA	CTTAAACAGCCTTGAAGAAA	AAAAAACAACCGCCCTGCCTTAATTA	3600
OY	3601	AATAGCCCACTTAAT	TGTTATATCAGCCTTCCCTTCTGCA	TTCAATTGAGAAATTCA	3660
Ds	3601	AATAGCCCACTTAAT	TGTTATATCAGCCTTCCCTTCTGCA	TTCAATTGAGAAATTCA	3660
OY	3661	AAGAAATTA	GACATCTCTACTGACCCAGAAACA	ATTATATCACTCTTCAGGCCCTGTG	3720
Ds	3661	AAGAAATTA	GACATCTCTACTGACCCAGAAACA	ATTATATCACTCTTCAGGCCCTGTG	3720
OY	3721	GGAGGCA	CAGTGGTAAAGGCTCTCTAACAGGTTTTTATATC	CCCTCCCTTAATACAAAT	3780
Ds	3721	GGAGGCA	CAGTGGTAAAGGCTCTCTAACAGGTTTTTATATC	CCCTCCCTTAATACAAAT	3780
OY	3781	GACAGAGTTTTGT	TATGGAACCTGGAATTTGCTGCTTCA	ATTCCTCCACCTGGCCTTAT	3840
Ds	3781	GACAGAGTTTTGT	TATGGAACCTGGAATTTGCTGCTTCA	ATTCCTCCACCTGGCCTTAT	3840
OY	3841	AGAGAAAC	TGAGATTGCTTCTGCAAAATTATG	GTCATGCAAAAGATTAATCTTAG	3900
Ds	3841	AGAGAAAC	TGAGATTGCTTCTGCAAAATTATG	GTCATGCAAAAGATTAATCTTAG	3900
OY	3901	ATTTTTTAT	TTATTTGGCAAAATATACAAAATGTCGAGAAATTA	AAATATCTGCTTAATCCAA	3960
Ds	3901	ATTTTTTAT	TTATTTGGCAAAATATACAAAATGTCGAGAAATTA	AAATATCTGCTTAATCCAA	3960
OY	3961	AGCTAAGTACTAAT	TTTGGTAAACACCAACTTGTGTAATATATATG	TAATAAGATCCATGA	4020
Ds	3961	AGCTAAGTACTAAT	TTTGGTAAACACCAACTTGTGTAATATATATG	TAATAAGATCCATGA	4020
OY	4021	ATTTCCCTTTT	AGTCANAGTGGGAAAGTGGATGTCCTTTT	TTTCTTATGTTACTTCA	4080
Ds	4021	ATTTCCCTTTT	AGTCANAGTGGGAAAGTGGATGTCCTTTT	TTTCTTATGTTACTTCA	4080
OY	4081	ATATAGAGAA	AAAGTAATGCTCAATATGCTGTTAAATTTAT	TTTAAATATATAGCTGAT	4140
Ds	4081	ATATAGAGAA	AAAGTAATGCTCAATATGCTGTTAAATTTAT	TTTAAATATATAGCTGAT	4140
OY	4141	CCGAGTCAG	TGGGTGTTACAACTACCTTGATATCA	ACAGATTCAGATTTCTTGTCTCT	4200
Ds	4141	CCGAGTCAG	TGGGTGTTGTTACAACTACCTTGATATCA	ACAGATTCAGATTTCTTGTCTCT	4200
OY	4201	TCTCCACTCC	CACTGCTTCACTTAACGCGCAAAAAGAAA	AAATTTTATATTA	4260
Ds	4201	TCTCCACTCC	CACTGCTTCACTTAACGCGCAAAAAGAAA	AAATTTTATATTA	4260
OY	4261	CTACTCA	AGACTAATATTTATTTATCTTATGTA	TTTATGCTGTTATTTATTTT	4320
Ds	4261	CTACTCA	AGACTAATATTTATTTATCTTATGTA	TTTATGCTGTTATTTATTTT	4320
OY	4321	ACTGTGTA	AAAAACAGGATTTGAGGGGAATACAGTTTAT	TTTATTTATTTATATAT	4380
Ds	4321	ACTGTGTA	AAAAACAGGATTTGAGGGGAATACAGTTTAT	TTTATTTATTTATATAT	4380
OY	4381	TTATTTAT	TTTATTTGGAATGGAATCTCTGTCA	CCACGCTGAGTGCAGTGTGCATC	4440
Ds	4381	TTATTTAT	TTTATTTGGAATGGAATCTCTGTCA	CCACGCTGAGTGCAGTGTGCATC	4440
OY	4441	TCAGATGA	CTGCAACCTCTGCTCTGAGTTCA	AGCAACTCTCCTGCGCCCTGCTT	4500
Ds	4441	TCAGATGA	CTGCAACCTCTGCTCTGAGTTCA	AGCAACTCTCCTGCGCCCTGCTT	4500
OY	4501	TACTTCTT	ATATCTGTTTAGTCAAGTGTACCTTA	CTTTTTCAAATGCTGAGAAATC	4560
Ds	4501	TACTTCTT	ATATCTGTTTAGTCAAGTGTGTACCTTA	CTTTTTCAAATGCTGAGAAATC	4560
OY	4561	TGCAATTA	AAAGACCACTTTTATTTTATTTAAGCTT	CTCATATCAATTTGGCCATGAT	4620
Ds	4561	TGCAATTA	AAAGACCACTTTTATTTTATTTAAGCTT	CTCATATCAATTTGGCCATGAT	4620
OY	4621	AACGTTTT	TCAGAGTGGCTCGAAACGGGGGAC	CCCTGGAACATATCTTGATATCATG	4680

Db	4621	AACTGTTTCAAGGTGGCTGGAA	CGGGGGACCCCTGGACATACCTTGATACATGGACA	4680
Qy	4681	CCATGGA	CACCTTCGATTCCTCTCTCTGAGTTCGACTCTTGATTTGTTCTGACAGACCTT	4740
Db	4681	CCATGGACACTTCGTGATCCTCTCTCTCTGAGTTTGACTTTGATTTGTTCTGACAGACCTT		4740
Qy	4741	TCGAGCCCGAAGTTTACACAGAAATTCATCTATCTTTCTCTACTTACTTAAAGTTTCT		4800
Db	4741	TCGAGCCCGAAGTTTACACAGAAATTCATCTATCTTTCTCTAGTACTTATATGTTTCT		4800
Qy	4801	TTTTCATTTAACTCTTTCATCTACTCGGAATTTATATGTAATATTCACAATCA	CCCCAGC	4860
Db	4801	TTTTCATTTAACTCTTTCATCTACTCGGAATTTATATGTAATATTCACAATCA	CCCCAGC	4860
Qy	4861	TCGATTTATAGATTTTCTTTCTCTGATAGTTGAAATGCTGGCAGATTAATATATAG		4920
Db	4861	TCGATTTATAGATTTTCTTTCTCTGATAGTTGAAATGCTGGCAGATTAATATATAG		4920
Qy	4921	ATATCAGAACTACTGGAATCTTCTCTGTCTAACTCTTTAAATAATCATGTTCCCTAAT		4980
Db	4921	ATATCAGAAATACTGGAATCTTCTCTGTCTTAATCTTTAAATAATCATGTTCCCTAAT		4980
Qy	4981	CTATCTTTCTTAATATTTGT	5000	
Db	4981	CTATCTTTCTTAATATTTGT	5000	

Db	241	CCCTGTCTATTAATAAATAATTAGTCGTGTTCCTTGGAACAAGCCCTGGAGTAGTACT	300
Qy	301	CGAAGAAGCTGAGGTGGAGGATCACTGAGCCCGAGAGGCTGAGGCTGCATGAACAAGTA	360
Db	301	CGAAGAAGCTGAGGTGGAGGATCACTGAGCCCGAGAGGCTGAGAGCTGCATGAACAAGTA	360
Qy	361	TCACCCGAGCTGGATTCGAGCCTGGAAGAACAAGAGAGAACCCCTGTTTCCAAAAA	420
Db	361	TCACCCGAGCTGGATTCGAGCCTGGAAGAACAAGAGAGAACCCCTGTTTCCAAAAA	420
Qy	421	AAAAAATAATGCAAGAAAGAATCATATAACTTGGACATTAACCTTTATGTGATG	480
Db	421	AAAAAATAATGCAAGAAAGAATCATATAACTTGGACATTAACCTTTATGTGATG	480
Qy	481	AAATTCACATCTTTTGAAGAATAATGAGATTCTGATATAAATGATTAATTAATTAT	540
Db	481	AAATTCACATCTTTTGAAGAATAATGAGATTCTGATATAAATGATTAATTAATTAT	540
Qy	541	ATTATTAATTCAAATGGAATTAATAATATTCTGAGAACTAGCTTCTCACTCTCAAGTTG	600
Db	541	ATTATTAATTCAAATGGAATTAATAATATTCTGAGAACTAGCTTCTCACTCTCAAGTTG	600
Qy	601	CAGTCAAAACTTTAATGATCTTTGGCCGGGTGGGGTGGCTCAAGCCGTGATCCAGAC	660
Db	601	CAGTCAAAACTTTAATGATCTTTGGCCGGGTGGGGTGGCTCAAGCCGTGATCCAGAC	660

RESULT 2

AX571875

LOCUS	AX571875	53332 bp	DNA	linear	PAT 29-NOV-2002
DEFINITION	Sequence 3 from Patent WO02061060.				

AX571875	AX571875.1	GI:26004000
ACCESSION		
VERSION		
NUMERIC		

KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1
Yan, C., Ketchum, K., di Francesco, V. and Beasley, E.M.

TITLE
Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof

JOURNAL
Patent: WO 02061060-A 3 08-AUG-2002;
PE Corporation (NY) (US)
Invention/Creation/Invention

```
FEATURES
SOURCE
LOCATION/QUALITY
1. 53332
/orcanism="Homo sapiens"
```

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

ORIGIN

Query Match	100.0%;	Score 5000;	DB 6;	Length 53332;
Best Local Similarity	100.0%;	Pred. No. 0;		

```

Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

Ph
0Y
1
1

TATAGGCGCAATGCCTGGCCACCGGCTATTCCCAGCATTTTGGGAGAGAGAGATATG 50
TAATGCGCAATGCCTGGCCACCGGCTATTCCCAGCATTTTGGGAGAGAGAGATATG 80

61 CTTGAGCTCAGGAA TTGAGACAGACCTACGTAA CATTAGTGAACCTCTGTCTGTACAAA 120

61 CTTGAGCTCAGSANTTGGAGACAGGCTTACGTACATGATGAACTCTGTCTGTACAA 120

121 TAATAAAGAAATTTTCCAGCATGCTGGCGGCACACCCCACTGCCAGCTATTTGGGAGGC 180

Db 121 TATATAAAGATTTTCAGGCATGTGGCGTCACCCCGAGTCCAGTATTGGGAGGC 180

181 TGAGTAGGAGGAATGCTTGAAGCCAGAGGTTGAAGACAAAGCCCTTAGGCACATAGTGAGA 240

Db 181 TGAGTAGAGGAAATGCTTGAAGCCAGAGTTGAAGCAAGCTTAGGCACATGTGACA 240

241 CCCTGTGTATATAAATAATATAGCTGGTGTCTTGGCACAGGCTGCAGCTAGCTACT 300

[illegible]

1321 GNGCTGTTAATAGTACTTTACATATAGTAAACATTAATTAATATATGTGAGAGAG 1380
QY 1381 CAGCAGAAAGAGAGAAAGTATTTATTCAAACTCTCTCTCATCACTATGGCTA 1440
QY 1381 CAGCAGAAAGAGAGAAAGTATTTATTCAAACTCTCTCTCATCACTATGGCTA 1440
QY 1441 AATATCATCTTTGTACAGTTAAGAACAAATAGGCTGACACAGATGTTTGAATAA 1500
Db 1441 AATATCATCTTTGTACAGTTAAGAACAAATAGGCTGACACAGATGTTTGAATAA 1500
QY 1501 TGAATGATAGGACACCTCTTAAGACTATTGATACACTATTGTTGAAGGCAAGAGAT 1560
Db 1501 TGAATGATAGGACACCTCTTAAGACTATTGATACACTATTGTTGAAGGCAAGAGAT 1560
QY 1561 GCAGTAGATATTTTCAACTTTTCTGTTTATGATCTGTGTTTCTTTGACTA 1620
Db 1561 GCAGTAGATATTTTCAACTTTTCTGTTTATGATCTGTGTTTCTTTGACTA 1620
QY 1621 AAGTTAGCTAGGTAGCAATTTGTTTAAAGCTGAAAACAAAATGCTTCAGATAAA 1680
Db 1621 AAGTTAGCTAGGTAGCAATTTGTTTAAAGCTGAAAACAAAATGCTTCAGATAAA 1680
QY 1681 GGTAGGAGAGAAAATCTCTCAACATCTCACTTTAGCAGCAGAAAACCTAATATCA 1740
Db 1681 GGTAGGAGAGAAAATCTCTCAACATCTCACTTTAGCAGCAGAAAACCTAATATCA 1740
QY 1741 TATCAACATCAATGATATCATTAATATCATTTGATAGATAGCAATGCAATCCCTA 1800
Db 1741 TATCAACATCAATGATATCATTAATATCATTTGATAGATAGCAATGCAATCCCTA 1800
QY 1801 AACTATGATATCAATAGCACTAATCTTGGCCAGAACAAACCTTAACTGTGCAAA 1860
Db 1801 AACTATGATATCAATAGCACTAATCTTGGCCAGAACAAACCTTAACTGTGCAAA 1860
QY 1861 TTTTATTTCTATTCAATTAACAGCTGCTCGTTTCACTGTGCACTGTAATGCAAGCA 1920
Db 1861 TTTTATTTCTATTCAATTAACAGCTGCTCGTTTCACTGTGCACTGTAATGCAAGCA 1920
QY 1921 TCCCTGTCTGATGTGAGTTTCTTGCACTGATTAAGAAAACGTGGAAGTTGTAGGCT 1980
Db 1921 TCCCTGTCTGATGTGAGTTTCTTGCACTGATTAAGAAAACGTGGAAGTTGTAGGCT 1980
QY 1981 GTTCAAGGAGAGCACTCATGTGAGTCATTAAGAGCTCAAGCTGTGCACTGTGCA 2040
Db 1981 GTTCAAGGAGAGCACTCATGTGAGTCATTAAGAGCTCAAGCTGTGCACTGTGCA 2040
QY 2041 AAGGAGAGAGCAAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 AAGGAGAGAGCAAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 2101 CTTTGAACCTAAGATTAATGAGTCAAGAGCTGTGCAAGAGCTGTGCAAGAGCTGTG 2160
Db 2101 CTTTGAACCTAAGATTAATGAGTCAAGAGCTGTGCAAGAGCTGTGCAAGAGCTGTG 2160
QY 2161 GGTGCTACATTTGTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2161 GGTGCTACATTTGTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 CACGAGAGAGCTGTTCAAGATATTGTTATCCAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 2221 CACGAGAGAGCTGTTCAAGATATTGTTATCCAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2281 AACCAAGAGAGTCAAGAGCTCTCTTTTGTAGCGGAGTGAATCTTCTGCGGTAGCT 2340
Db 2281 AACCAAGAGAGTCAAGAGCTCTCTCTTTTGTAGCGGAGTGAATCTTCTGCGGTAGCT 2340
QY 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
Db 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
QY 2401 TTTGGGGAATTTCTTTGTAACCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
Db 2401 TTTGGGGAATTTCTTTGTAACCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460

2461 ACTCACTTGCTTGAAGTGGAGAAAGTACTTCTCAAAATCTTAAGGCTCTTAATTA 2520
Db 2461 ACTCACTTGCTTGAAGTGGAGAAAGTACTTCTCAAAATCTTAAGGCTCTTAATTA 2520
QY 2521 AAGGCTGAACTTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 2521 AAGGCTGAACTTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
QY 2581 TAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
Db 2581 TAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
QY 2641 AGTGAATGAG 2700
Db 2641 AGTGAATGAG 2700
QY 2701 AAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
Db 2701 AAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
QY 2761 TTTCAAGAGTGTATCAATTAATGATCTTTTCAAGGAGCTTCAAGAAAGTCCCTT 2820
Db 2761 TTTCAAGAGTGTATCAATTAATGATCTTTTCAAGGAGCTTCAAGAAAGTCCCTT 2820
QY 2821 GGGGTACAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Db 2821 GGGGTACAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
QY 2881 GGTTTTACAGAGAGCTGATCTTTCTTTGCAATTAATTAATTAATTAATTAATTA 2940
Db 2881 GGTTTTACAGAGAGCTGATCTTTCTTTGCAATTAATTAATTAATTAATTAATTA 2940
QY 2941 CTCTAGCTAACAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Db 2941 CTCTAGCTAACAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
QY 3001 CTTCAGAGCTGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 3001 CTTCAGAGCTGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
QY 3061 TTTCAAG 3120
Db 3061 TTTCAAG 3120
QY 3121 AAGAGAGATTAATGAG 3180
Db 3121 AAGAGAGATTAATGAG 3180
QY 3181 GAGCAGAGCTTCCAGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Db 3181 GAGCAGAGCTTCCAGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
QY 3241 CAACAG 3300
Db 3241 CAACAG 3300
QY 3301 CTTAG 3360
Db 3301 CTTAG 3360
QY 3361 GGTGGGTGAAGCTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
Db 3361 GGTGGGTGAAGCTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
QY 3421 CTGGAG 3480
Db 3421 CTGGAG 3480
QY 3481 GAGAGATCAATGAG 3540
Db 3481 GAGAGATCAATGAG 3540

QY 3541 GAGAGCATGATTAATTAAACAGCTTGAAGAAAAACAACCTGCTGCTTAATTA 3600
 DB 3541 GAGAGCATGATTAATTAAACAGCTTGAAGAAAAACAACCTGCTGCTTAATTA 3600
 QY 3601 AATCAGCCCACTTAATGTTTATCAGCCCTTCCCTTGGCAATTAATCAGAGATTA 3660
 DB 3601 AATCAGCCCACTTAATGTTTATCAGCCCTTCCCTTGGCAATTAATCAGAGATTA 3660
 QY 3661 AATCAGCCCACTTAATGTTTATCAGCCCTTCCCTTGGCAATTAATCAGAGATTA 3660
 DB 3661 AATCAGCCCACTTAATGTTTATCAGCCCTTCCCTTGGCAATTAATCAGAGATTA 3660
 QY 3721 GAGGCAAGTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAGAT 3780
 DB 3721 GAGGCAAGTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAGAT 3780
 QY 3781 GAGGCAAGTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAGAT 3780
 DB 3781 GAGGCAAGTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAGAT 3780
 QY 3841 AGAAGAACTGAAGTGGTCTTGAAGATTAATGATCAATGCAAAAGATGATTAATCAG 3900
 DB 3841 AGAAGAACTGAAGTGGTCTTGAAGATTAATGATCAATGCAAAAGATGATTAATCAG 3900
 QY 3901 AATTTTATATTTGCAAAATACCAAAATGCTGAGAAATTAATATCTTATCCAA 3960
 DB 3901 AATTTTATATTTGCAAAATACCAAAATGCTGAGAAATTAATATCTTATCCAA 3960
 QY 3961 AGCTAAGTAAATTTTGTAAACCAACCACTTGTAAATATATGTAATTAATAGCTGA 4020
 DB 3961 AGCTAAGTAAATTTTGTAAACCAACCACTTGTAAATATATGTAATTAATAGCTGA 4020
 QY 4021 AATCCCTTTTGAAGTGGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4080
 DB 4021 AATCCCTTTTGAAGTGGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4080
 QY 4081 ATGAGAGAAAGTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4140
 DB 4081 ATGAGAGAAAGTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4140
 QY 4141 CCAGTGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4200
 DB 4141 CCAGTGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4200
 QY 4201 TCTCCACTCCCACTGCTTCACTTAATCTGCGCCAAAAGAAAAAGAAAAATTTTATTA 4260
 DB 4201 TCTCCACTCCCACTGCTTCACTTAATCTGCGCCAAAAGAAAAAGAAAAATTTTATTA 4260
 QY 4261 CTACTACAAGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320
 DB 4261 CTACTACAAGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4320
 QY 4321 ACTTGTAAAAACAGATTTGAGGGGACATACAGTTTTTATTTTATTTATTTATTTAT 4380
 DB 4321 ACTTGTAAAAACAGATTTGAGGGGACATACAGTTTTTATTTTATTTATTTATTTATTTAT 4380
 QY 4381 TTATTTATTTATTTGGAATGGAATCTCTGTCAACCGCGTGGAGTGCAGTGGGATC 4440
 DB 4381 TTATTTATTTATTTGGAATGGAATCTCTGTCAACCGCGTGGAGTGCAGTGGGATC 4440
 QY 4441 TCAATGACGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
 DB 4441 TCAATGACGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
 QY 4501 TACTTTCTTAATCTGTTTATGCTAATGCTAATCTTAACTTTTCAATGCTGAGAACATC 4560
 DB 4501 TACTTTCTTAATCTGTTTATGCTAATGCTAATCTTAACTTTTCAATGCTGAGAACATC 4560
 QY 4561 TGCATTAAGGACCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4620
 DB 4561 TGCATTAAGGACCACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4620
 QY 4621 AATGTTTTTCAAGTGGCTCGGAACGGGGGCACTCGGAACATCTTGAATACATGGGCA 4680

DB 4621 AATGTTTTTCAAGTGGCTCGGAACGGGGGCACTCGGAACATCTTGAATACATGGGCA 4680
 QY 4681 CCATGACACTTCTGATCTCTCTCTGAGTCTGACTTGAATGTTCTGACAGACTT 4740
 DB 4681 CCATGACACTTCTGATCTCTCTCTGAGTCTGACTTGAATGTTCTGACAGACTT 4740
 QY 4741 TCCAGCCGGAAGTTACACAAATTCATCTTTCTTCTAGTTCCTTATGTTTCT 4800
 DB 4741 TCCAGCCGGAAGTTACACAAATTCATCTTTCTTCTAGTTCCTTATGTTTCT 4800
 QY 4801 TTTTCAATTAATCTTTTCACTGAGGAAATTAATTAATTAATTAATTAATTAATTAATTA 4860
 DB 4801 TTTTCAATTAATCTTTTCACTGAGGAAATTAATTAATTAATTAATTAATTAATTAATTA 4860
 QY 4861 TCCATTTATAGATTTCTTTTCTGATGTTGAATGCTGCAATGATTAATTAATTAATTA 4920
 DB 4861 TCCATTTATAGATTTCTTTTCTGATGTTGAATGCTGCAATGATTAATTAATTAATTA 4920
 QY 4921 ATCTCAGAAATCTTGAATCTTTCTGATGTTGAATGCTGCAATGATTAATTAATTA 4980
 DB 4921 ATCTCAGAAATCTTGAATCTTTCTGATGTTGAATGCTGCAATGATTAATTAATTA 4980
 QY 4981 CTATCTTTCTTAATTTGT 5000
 DB 4981 CTATCTTTCTTAATTTGT 5000

RESULT 3
 AC007242/c 112748 bp DNA linear PRI 30-SEP-2000
 LOCUS Homo sapiens BAC clone RP11-182H9 from 2, complete sequence.
 DEFINITION AC007242.3 GI:5931465
 ACCESSION AC007242.3 GI:5931465
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 112748)
 Suleston,J.E. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE
 PUBMED
 9847074
 REFERENCE
 2 (bases 1 to 112748)
 Kozlowicz,A., Stoneking,T., Hawkins,M. and Hawryshko,C.
 The sequence of Homo sapiens BAC clone RP11-182H9
 Unpublished
 JOURNAL
 3 (bases 1 to 112748)
 REFERENCE
 Waterston,R.H.
 Direct Submission
 Submitted (07-APR-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 112748)
 REFERENCE
 Waterston,R.H.
 Direct Submission
 Submitted (28-SEP-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 112748)
 REFERENCE
 Waterston,R.
 Direct Submission
 Submitted (30-SEP-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 28, 1999 this sequence version replaced gi:5001521.
 COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sepiens@wustl.edu
 ----- Summary Statistics

Center project name: H_NH0182H09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moom, P.Y., Zhao, B., Frengen, B., Tateo, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-53H10, 200 bp overlap; the clone sequenced to the right is RP11-309N8. Actual start of this clone is at base position 100126 of RP11-53H10; actual end is at 112748 of RP11-182H9.

The reads used to determine the sequence at the region of base positions 31840 to 112748 consists of numerous polymorphic base changes represented from both NH0182H09 and RP11-309N8. It is unknown which base calls belong to which clone.

FEATURES

Source

1. 112748
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-182H9"
/clone_11b="RPCI-11"
1033. 1331
/rpt_family="Alu"
3574. 3737
/rpt_family="MIR"
3814. 4009
/rpt_family="MIR1_type"
4048. 4358
/rpt_family="Alu"
6038. 6457
/rpt_family="MALR"
6477. 6680
/rpt_family="L2"
6681. 6978
/rpt_family="Alu"
7386. 7542
/rpt_family="Alu"
7543. 7567
/rpt_family="TTTG)n"
7568. 7702
/rpt_family="Alu"
8575. 8596
/rpt_family="AT-rich"

repeat_region 8599. 9529
/rpt_family="L1"
repeat_region 9685. 9875
/rpt_family="MIR"
repeat_region 10125. 10238
/rpt_family="MIR"
repeat_region 10357. 10458
/rpt_family="MIR1_type"
repeat_region 10484. 10784
/rpt_family="Alu"
repeat_region 10898. 11020
/rpt_family="MIR"
STS 11835. 12166
/db_xref="GI:1234078"
repeat_region 12082. 12118
/rpt_family="(CA)n"
repeat_region 12210. 12521
/rpt_family="Alu"
repeat_region 13068. 13394
/rpt_family="L2"
repeat_region 13429. 13479
/rpt_family="Alu"
repeat_region 13511. 13812
/rpt_family="Alu"
repeat_region 13967. 14092
/rpt_family="L2"
repeat_region 14111. 14139
/rpt_family="AT-rich"
repeat_region 14206. 14515
/rpt_family="Alu"
repeat_region 14516. 14559
/rpt_family="Alu"
repeat_region 14723. 14767
/rpt_family="AT-rich"
repeat_region 15360. 15575
/rpt_family="MIR"
repeat_region 15691. 15734
/rpt_family="L2"
repeat_region 16277. 16577
/rpt_family="Alu"
repeat_region 16608. 16649
/rpt_family="L2"
repeat_region 16793. 16904
/rpt_family="MIR"
repeat_region 16919. 17003
/rpt_family="MIR"
repeat_region 19659. 19795
/rpt_family="GA-rich"
repeat_region 20648. 20707
/rpt_family="L2"
repeat_region 20881. 21187
/rpt_family="Alu"
repeat_region 22002. 22304
/rpt_family="Alu"
repeat_region 23542. 23599
/rpt_family="L2"
repeat_region 23944. 24249
/rpt_family="Alu"
repeat_region 25966. 26170
/rpt_family="L1"
repeat_region 26171. 26461
/rpt_family="Alu"
repeat_region 26476. 26808
/rpt_family="MALR"
repeat_region 26809. 27119
/rpt_family="Alu"
repeat_region 27120. 27185
/rpt_family="MALR"
repeat_region 27186. 27530
/rpt_family="L1"
repeat_region 27858. 27971
/rpt_family="Alu"
repeat_region 28007. 28310

```
repeat_region /rpt_family="Alu"  
28872..29076  
repeat_region /rpt_family="MIR"  
29419..29758  
repeat_region /rpt_family="L1"  
30018..30265  
misc_feature /rpt_family="Alu"  
30648..30739  
/note="match to EST AA436054 (NID:92140968) zu01c12.r1"  
repeat_region /rpt_family="MER1_type"  
31066..31159  
Query Match 99.5%; Score 4973.4; DB 9; Length 112748;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4999; Conservative 0; Mismatches 1; Indels 5; Gaps 2;  
Qy 1 TATAGCCAAATGCTGTGGCTCAGCGCTGTTCCAGCACTTTGGAGGACAGAGGATCG 60  
Db 73582 TATAGGCCAATGCTGTGGCTCAGCGCTGTTCCAGCACTTTGGAGGACAGAGGATCG 73523  
Qy 61 CTGAGCTCAGGAATGGAGACAAGCCTAGTAACATAGTAACCTCTGTCTACAA 120  
Db 73522 CTGAGCTCAGGAATGGAGACAAGCCTAGTAACATAGTAACCTCTGTCTACAA 73463  
Qy 121 TAAATAAAGATTTTCCAGGCATGCTGCGTGCACCCCACTGCTGCTGTTGGAGGC 180  
Db 73462 TAAATAAAGATTTTCCAGGCATGCTGCGTGCACCCCACTGCTGCTGTTGGAGGC 73403  
Qy 181 TGAGGTGAGAGAAATGCTTTGAAGCCAGAGATTGAAGCAAGCCTAGGCAACATAGTGA 240  
Db 73402 TGAGGTGAGAGAAATGCTTTGAAGCCAGAGATTGAAGCAAGCCTAGGCAACATAGTGA 73343  
Qy 241 CCTGTGCTATATAAATAATATAGTGTGTCTTGGCA CAGGCTGCAAGCTGACT 300  
Db 73342 CCTGTGCTATATAAATAATATAGTGTGTCTTGGCA CAGGCTGCAAGCTGACT 73283  
Qy 301 CGAAGACTGAGGTGGAGATCACTGAGCCAGAGAGCTGAGAGCTGCACTGGAACAGTGA 360  
Db 73282 CGAAGACTGAGGTGGAGATCACTGAGCCAGAGAGCTGAGAGCTGCACTGGAACAGTGA 73223  
Qy 361 TCACCCAGCTGAGTTCCAGCCTGGAGACAGAGGAGACCTGTTTCAAAAAA 420  
Db 73222 TCACCCAGCTGAGTTCCAGCCTGGAGACAGAGGAGACCTGTTTCAAAAAA 73163  
Qy 421 AAAAAAATGCAAGAAAAAATCATTAACCTTGAACCTGGACATTAATTATATATAT 480  
Db 73162 AAAAAAATGCAAGAAAAAATCATTAACCTTGAACCTGGACATTAATTATATATAT 73103  
Qy 481 AATTCACAATCTTTAGGAAGAAATTAGCATTTCTGATAAATGTAATTAATATAT 540  
Db 73102 AATTCACAATCTTTAGGAAGAAATTAGCATTTCTGATAAATGTAATTAATATAT 73043  
Qy 541 ATTATTAATTCAAATGAAATTAATATATCTGAGAACTAGCTTCACTCTCAATTT 600  
Db 73042 ATTATTAATTCAAATGAAATTAATATATCTGAGAACTAGCTTCACTCTCAATTT 72983  
Qy 601 CAGTCAAACTTTAATGCTCTTGGCGGGGTGCGGTGCTCAACCTGTAATCCAGAC 660  
Db 72982 CAGTCAAACTTTAATGCTCTTGGCGGGGTGCGGTGCTCAACCTGTAATCCAGAC 72923  
Qy 661 TTTGGAGGCGAGAGCGGGTGAATCAAGGTTAGAGATGAGACATTCCTGGTAA 720  
Db 72922 TTTGGAGGCGAGAGCGGGTGAATCAAGGTTAGAGATGAGACATTCCTGGTAA 72863  
Qy 721 CCGTGAACCTGCTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 780  
Db 72862 CCGTGAACCTGCTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 72803  
Qy 781 TGTAGTCCAGAGCTGCTCAGAGGCTGAGGCAAGAAATGAGTGAACCCGGAGGCGAG 840  
Db 72802 TGTAGTCCAGAGCTGCTCAGAGGCTGAGGCAAGAAATGAGTGAACCCGGAGGCGAG 72743  
Qy 841 CTTCAGTGAAGCGAGATTGGCGCACTGCACTCAGCTGGCGACAGTGCAGACTCTG 900
```

```
Db 72742 CTTCAGTGAAGCGAGATTGGCGCACTGCACTCAGCTGGCGACAGTGCAGACTCTG 72683  
Qy 901 TCTC-AAAAAAAAAAAAAAAAAGTGAATGCTTTGAGCCAAATGCTTCTCTTT 959  
Db 72682 TCTCAAAAAAAAAAAAAAAAAAGTGAATGCTTTGAGCCAAATGCTTCTCTTT 72623  
Qy 960 CTCTCTCTCTTTTTTTTTTTTTTCAAAAAAATCTGTAATGTAATCTGGAAATGGC 1019  
Db 72622 CTCTCTCTCTTTTTTTTTTTTTTCAAAAAAATCTGTAATGTAATCTGGAAATGGC 72563  
Qy 1020 TTAAGTCTCTCTCTGTCGCAATTTGAAATGAAAAAATATAGTCTATTAATTA 1079  
Db 72562 TTAAGTCTCTCTCTGTCGCAATTTGAAATGAAAAAATATAGTCTATTAATTA 72503  
Qy 1080 CTTGAACATTTTAAAAAACAATCATGAGGTTCAAAATTCATAATCAATTAATTTG 1139  
Db 72502 CTTGAACATTTTAAAAAACAATCATGAGGTTCAAAATTCATAATCAATTAATTTG 72443  
Qy 1140 TGAATATGACATTAATCTTATTTTCCCTTAATAATGATTTTATATATCTCCAT 1199  
Db 72442 TGAATATGACATTAATCTTATTTTCCCTTAATAATGATTTTATATATCTCCAT 72383  
Qy 1200 CTGCTCACTTATGATTAATTAATTAATGAGCAATTAATCTTGAATCTAACAGAA 1259  
Db 72382 CTGCTCACTTATGATTAATTAATTAATGAGCAATTAATCTTGAATCTAACAGAA 72323  
Qy 1260 AAGTGTGCATTTGAAGATCAACAGCAGAAACCAATTTAAGCAGATCTTGAACAT 1319  
Db 72322 AAGTGTGCATTTGAAGATCAACAGCAGAAACCAATTTAAGCAGATCTTGAACAT 72263  
Qy 1330 TGTGCTGTAATATAGTCTTATTAATATAGTAACATTAATTAATATATGAGAA 1379  
Db 72262 TGTGCTGTAATATAGTCTTATTAATATAGTAACATTAATTAATATATGAGAA 72203  
Qy 1380 GCAGCAAGAAAGAAAGATTTTCAATCAACCTCTCTCTCCATCAACATTTGCT 1439  
Db 72202 GCAGCAAGAAAGAAAGATTTTCAATCAACCTCTCTCTCCATCAACATTTGCT 72143  
Qy 1440 AATATCATCTTTGAGAGTTAAGAAACAATAGTGTCTCAACATATGTTTGAATTA 1499  
Db 72142 AATATCATCTTTGAGAGTTAAGAAACAATAGTGTCTCAACATATGTTTGAATTA 72083  
Qy 1500 ATGAATGAATGCAACCTCTTGAACATTAATGATACATTAATGTTTGAATTA 1559  
Db 72082 ATGAATGAATGCAACCTCTTGAACATTAATGATACATTAATGTTTGAATTA 72023  
Qy 1560 TGCAGTGAATTTTCACTTTTCCGTTTATGATTTCTGAGTTCTTGAATTA 1619  
Db 72022 TGCAGTGAATTTTCACTTTTCCGTTTATGATTTCTGAGTTCTTGAATTA 71963  
Qy 1620 AAGTGAAGTGAACAAATTTGTTTAAAGTGAACCAAAATGCTTCAATTA 1679  
Db 71962 AAGTGAAGTGAACAAATTTGTTTAAAGTGAACCAAAATGCTTCAATTA 71903  
Qy 1680 AGGTAGGAGAAATAATCTCTCAACATGCTTGAACCTTGAACCAAAATGCTTCA 1739  
Db 71902 AGGTAGGAGAAATAATCTCTCAACATGCTTGAACCTTGAACCAAAATGCTTCA 71843  
Qy 1740 ATATCAACATCAATGATATCATATAATATCAATGATAGTAATGCAATCCCTA 1799  
Db 71842 ATATCAACATCAATGATATCATATAATATCAATGATAGTAATGCAATCCCTA 71783  
Qy 1800 AAAAATATGATATCAATAGCACTAACTTGGCCAGAAAGAACTTAACTTGGCCAA 1859  
Db 71782 AAAAATATGATATCAATAGCACTAACTTGGCCAGAAAGAACTTAACTTGGCCAA 71723  
Qy 1860 ATTTATATCTATTAATTAACAGTGTCTGTTTGAAGTGTGCAATCTGAATGAAGA 1919  
Db 71722 ATTTATATCTATTAATTAACAGTGTCTGTTTGAAGTGTGCAATCTGAATGAAGA 71663  
Qy 1920 ATCCGTCTGATGAGATTTCTTCACTGATTAAGAAAACTGTGAATTTGAGGC 1979
```

```
Db 71662 ATCCCTGCTGATGTGAGATTCTTGCACTGATTAAGGAAAACTGCTGAAGTTGTAGGC 71603
Qy 1980 TGTCTCAGGCAAGGCATCATGTGATGATCATATGAAAGCTCAAGCTGCTGACCTCTGGCA 2039
Db 71602 TGCTCAAGGCAAGGCATCATGTGATGATCATATGAAAGCTCAAGCTGCTGACCTCTGGCA 71543
Qy 2040 AAAAGGAGAGAAACAAGATAGAGAGGCAAGTGGGGAAGGTTCAAGTGGGGTTTCT 2099
Db 71542 AAAAGGAGAGAAACAAGATAGAGAGGCAAGTGGGGAAGGTTCAAGTGGGGTTTCT 71483
Qy 2100 CCTGAACCTACAAAGATTATGGGTCAAGAGCTGTGCAAAAGCTGTACAGCTGGATGC 2159
Db 71482 CCTGAACCTACAAAGATTATGGGTCAAGAGCTGTGCAAAAGCTGTACAGCTGGATGC 71423
Qy 2160 AGCTGCTACATTTGTTCAAGAGGAGGCAAGCAAGCTGTGCAAGAGAGTCAAGCTGAG 2219
Db 71422 AGCTGCTACATTTGTTCAAGAGGAGGCAAGCAAGCTGTGCAAGAGTCAAGCTGAG 71363
Qy 2220 ACCAGAGAGGCTGCTCAAGGATTTGTATCCCAAGAGAGAGCATCTTCTCTATGAT 2279
Db 71362 ACCAGAGAGGCTGCTCAAGGATTTGTATCCCAAGAGAGAGCATCTTCTCTATGAT 71303
Qy 2280 AAACCAAGAGTTCAGACACTCCCTTTTGTAGCGGAGCTGATTTCTTGCGGATAGTC 2339
Db 71302 AAACCAAGAGTTCAGACACTCCCTTTTGTAGCGGAGCTGATTTCTTGCGGATAGTC 71243
Qy 2340 TAAACCAATAAATATGAAAAATTTCTATTAAGTCAAGAAAAATTTATGCTGATATCA 2399
Db 71242 TAAACCAATAAATATGAAAAATTTCTATTAAGTCAAGAAAAATTTATGCTGATATCA 71183
Qy 2400 ATTTGGGGAATTTCTTTGTAACCAAAAGGAAAAATATCTCTGCTTGGCTGGCTGACA 2459
Db 71182 ATTTGGGGAATTTCTTTGTAACCAAAAGGAAAAATATCTCTGCTTGGCTGGCTGACA 71123
Qy 2460 AACTCACTTGGCTTGAAGTCGAGAAAGTAGTCTCTCAAAATCTCTAAGGCTCTAAATTA 2519
Db 71122 AACTCACTTGGCTTGAAGTCGAGAAAGTAGTCTCTCAAAATCTCTAAGGCTCTAAATTA 71063
Qy 2520 CAGAGCTGAAACTTAAAGGCAAGCTGCAAGTATTAAGTGTGATGATGATTTGAACT 2579
Db 71062 CAGAGCTGAAACTTAAAGGCAAGCTGCAAGTATTAAGTGTGATGATGATTTGAACT 71003
Qy 2580 TTAGTAATTTAGTTCATGATTTATTAAGCATGATGATTTATCCCTTACAGCAATTAAT 2639
Db 71002 TTAGTAATTTAGTTCATGATTTATTAAGCATGATGATTTATCCCTTACAGCAATTAAT 70943
Qy 2640 AAGTGAACATGAAAAAAGGCAAGCTTAAACAGAAAAAGTTGCAAAACATCATCA 2699
Db 70942 AAGTGAACATGAAAAAAGGCAAGCTTAAACAGAAAAAGTTGCAAAACATCATCA 70883
Qy 2700 AAGAGATTTAGTTAACTGTAATGTTAAAGACATTTTATAGTGAAGAAAGATGAT 2759
Db 70882 AAGAGATTTAGTTAACTGTAATGTTAAAGACATTTTATAGTGAAGAAAGATGAT 70823
Qy 2760 ATTTGAGAGATTGATACCATTTATGCTTTTCAAGGATCTTTCAAGAAAGTGCCTTT 2819
Db 70822 ATTTGAGAGATTGATACCATTTATGCTTTTCAAGGATCTTTCAAGAAAGTGCCTTT 70763
Qy 2820 GGGGGTACAGAAAGCTTGAAGAAACATTTGAAGAGTGAAGTGAAGCAATTAAGAAAAA 2879
Db 70762 GGGGGTACAGAAAGCTTGAAGAAACATTTGAAGAGTGAAGTGAAGCAATTAAGAAAAA 70703
Qy 2880 TGGTTTTACAGGCACTGAAATCTTTATCTTTGATTAATTTATTTCTGCTCTTTCTTT 2939
Db 70702 TGGTTTTACAGGCACTGAAATCTTTATCTTTGATTAATTTATTTCTGCTCTTTCTTT 70643
Qy 2940 TCTCTAGCTACAGACCTTAAAGAAAGCATGTTCCATGATCTTCAATTTCAACCCAGGG 2999
Db 70642 TCTCTAGCTACAGACCTTAAAGAAAGCATGTTCCATGATCTTCAATTTCAACCCAGGG 70583
Qy 3000 ACTTCAAGCTGCGCTGCGCAGAGATTCAAGAGTAAAGGCAAGGAGTAAAGATGATG 3059
Db 70582 ACTTCAAGCTGCGCTGCGCAGAGATTCAAGAGTAAAGGCAAGGAGTAAAGATGATG 70523

Qy 3060 TTTTCAGAGAGAGATCTGAGGCAAGGTTTTCATGTTGGTGAAGTCAAGCTGATGAT 3119
Db 70522 TTTTCAGAGAGAGATCTGAGGCAAGGTTTTCATGTTGGTGAAGTCAAGCTGATGAT 70463
Qy 3120 CAAGAAGAAATTTAATGATGATGTTCTGACAGAGGCGGCGCTTGGCTTCAAGGCAATTA 3179
Db 70462 CAAGAAGAAATTTAATGATGATGTTCTGACAGAGGCGGCGCTTGGCTTCAAGGCAATTA 70403
Qy 3180 TGAGCGAGCTTCCCAAGCTGCTCTGCAATGCTGTCAATTTTCCCTGGGAAAAAAG 3239
Db 70402 TGAGCGAGCTTCCCAAGCTGCTCTGCAATGCTGTCAATTTTCCCTGGGAAAAAAG 70343
Qy 3240 TCAACACTAATAAAGATGTTCTTTCTCTCTTCCCTTCAACCGCTCTTTTCCCATTC 3299
Db 70342 TCAACACTAATAAAGATGTTCTTTCTCTCTTCCCTTCAACCGCTCTTTTCCCATTC 70283
Qy 3300 CCTTGAAGAGAGAGAGAGGCTCCCTTTTGGGAGAGGCTCATCTTACTTGAACCTTGAGA 3359
Db 70282 CCTTGAAGAGAGAGAGAGGCTCCCTTTTGGGAGAGGCTCATCTTACTTGAACCTTGAGA 70223
Qy 3360 AGCTGGTGAAGGCTCTTATGCAAGATTAAAGGAGATTAGCAGTCAAGTACACATA 3419
Db 70222 AGCTGGTGAAGGCTCTTATGCAAGATTAAAGGAGATTAGCAGTCAAGTACACATA 70163
Qy 3420 GCTGGAGAGACTTGAAGATGAGAGTCCCGCCCGCCCAATTTCAATTAATAAGCCAG 3479
Db 70162 GCTGGAGAGACTTGAAGATGAGAGTCCCGCCCGCCCAATTTCAATTAATAAGCCAG 70103
Qy 3480 TGAGACATCATAGAATTCATAGCACTGAGACCTGTGCAAGACACANTGGCCGACAGG 3539
Db 70102 TGAGACATCATAGAATTCATAGCACTGAGACCTGTGCAAGACACANTGGCCGACAGG 70043
Qy 3540 AGAGAGACATGATTAATTAACAGGCTTGAAGAAAAAACAACCTGCGCTGCTTAATTA 3599
Db 70042 AGAGAGACATGATTAATTAACAGGCTTGAAGAAAAAACAACCTGCGCTGCTTAATTA 69983
Qy 3600 AAATCAGCCCACTTAATATGTTTATCAGGCTTCCCTTCTGCAATTAATTAAGAAATTC 3659
Db 69982 AAATCAGCCCACTTAATATGTTTATCAGGCTTCCCTTCTGCAATTAATTAAGAAATTC 69923
Qy 3660 AAAGAAATTAAGACATCTCTACTGACCTGAAAGAAACAAATTAATGACTTCAAGGCT 3719
Db 69922 AAAGAAATTAAGACATCTCTACTGACCTGAAAGAAACAAATTAATGACTTCAAGGCT 69863
Qy 3720 GGGAGGCAAGTGGTAAAGGCTCTTAACAGATTTTAAATCCCTCCCTAAACACA 3779
Db 69862 GGGAGGCAAGTGGTAAAGGCTCTTAACAGATTTTAAATCCCTCCCTAAACACA 69803
Qy 3780 TGACAGATTTTGAATGCAACCTGGAATTTGCTGCTCAATTCCTGCAAGCTGGCTTTA 3839
Db 69802 TGACAGATTTTGAATGCAACCTGGAATTTGCTGCTCAATTCCTGCAAGCTGGCTTTA 69743
Qy 3840 TGAAGAAATCGAAGTGGTTTCTGCAAAATTAATGTTAATGCAAGAAATGATTAATCTTA 3899
Db 69742 TGAAGAAATCGAAGTGGTTTCTGCAAAATTAATGTTAATGCAAGAAATGATTAATCTTA 69683
Qy 3900 GATTTTATATTTGCAAAATACAAATATGCTGGAATTAATAATCTGTTATCCA 3959
Db 69682 GATTTTATATTTGCAAAATACAAATATGCTGGAATTAATAATCTGTTATCCA 69623
Qy 3960 AAGCTAAGTAAATTTTGGTAAACAACTTTGTTAAATATATGTAAGATCAATG 4019
Db 69622 AAGCTAAGTAAATTTTGGTAAACAACTTTGTTAAATATATGTAAGATCAATG 69563
Qy 4020 AATTCCTCTTTAAGTCAAGGAGGGAAGTGGATGCTGCTTTTCTTATGTAATCC 4079
Db 69562 AATTCCTCTTTAAGTCAAGGAGGGAAGTGGATGCTGCTTTTCTTATGTAATCC 69503
Qy 4080 AATGAGAGAAAGTAAATGCTCAATGCTTAAATATTAATTTAAATATATAGCTGA 4139
Db 69502 AATGAGAGAAAGTAAATGCTCAATGCTTAAATATTAATTTAAATATATAGCTGA 69443
```

```

QY      4140 TCCGAGTGCAGTGTGTTTACAACTATCTGATCACAACAGTTACAGATTCTTGTGTC 4199
      |||
      69442 TCCGAGTGCAGTGTGTTTACAACTATCTGATCACAACAGTTACAGATTCTTGTGTC 69383
QY      4200 TTCTCCACTCCCACTGCTTCACTTAACTGGCCAAAAACGAAAAAAGAAAAATTTTATATA 4259
      |||
      69382 TTCTCCACTCCCACTGCTTCACTTAACTGGCCAAAAACGAAAAAAGAAAAATTTTATATA 69323
QY      4260 ACTACTACAAAGACTAAATATTTATTTATTTATCTAGTATTTATGCTGTTATTTATTTT 4319
      |||
      69322 ACTACTACAAAGACTAAATATTTATTTATTTATCTAGTATTTATGCTGTTATTTATTTT 69263
QY      4320 TACTGTGTTAAAACGAGATTGTAGGGGACATACAGTTTATTTTATTTATTTATTA-- 4377
      |||
      69262 TACTGTGTTAAAACGAGATTGTAGGGGACATACAGTTTATTTTATTTATTTATTTATTA 69203
QY      4378 --TATTTATTTATTTATTTTGGAAATGGAATCTGTGACCCGCGTGGAGTCAAGTGG 4435
      |||
      69202 TTTATTTATTTATTTATTTTGGAAATGGAATCTGTGACCCGCGTGGAGTCAAGTGG 69143
QY      4436 CGATCTCAGTGCAGTGCACACTGCGCTGCTGAGTTCAGACAACTCTGCGCCCTGCGCC 4495
      |||
      69142 CGATCTCAGTGCAGTGCACACTGCGCTGCTGAGTTCAGACAACTCTGCGCCCTGCGCC 69083
QY      4496 CTTTATATCTTTCTTAAATCTGTTTATGTCATGCTGTAACCTTAACTTTTTCATGCTGAGA 4555
      |||
      69082 CTTTATATCTTTCTTAAATCTGTTTATGTCATGCTGTAACCTTAACTTTTTCATGCTGAGA 69023
QY      4556 ACATCTGCATTAAGGACCACTTTTATTTATTTTCTTAAAGCTTCTCATATCAATTTGGCC 4615
      |||
      69022 ACATCTGCATTAAGGACCACTTTTATTTATTTTCTTAAAGCTTCTCATATCAATTTGGCC 68963
QY      4616 ATGTAACTGTTTTCAGAGTGTGCTCGGAAACGGGGGACCCCTGGAACATATCTTGATATAT 4675
      |||
      68962 ATGTAACTGTTTTCAGAGTGTGCTCGGAAACGGGGGACCCCTGGAACATATCTTGATATAT 68903
QY      4676 GGGGACCATGGAACATCTTCTGATCTCTCTCTGATGTTTGAATTTGTTTGGACAG 4735
      |||
      68902 GGGGACCATGGAACATCTTCTGATCTCTCTCTGATGTTTGAATTTGTTTGGACAG 68843
QY      4736 ACCCTTCCAGCCGGAAGTTTACACAGAAATCACTTATCTTTCTTCTAGTTACTTATATG 4795
      |||
      68842 ACCCTTCCAGCCGGAAGTTTACACAGAAATCACTTATCTTTCTTCTAGTTACTTATATG 68783
QY      4796 TTTCTTTTCAATTAACCTTCTTCACTGACGCGGAATTTATTTGATTTACACATCAC 4855
      |||
      68782 TTTCTTTTCAATTAACCTTCTTCACTGACGCGGAATTTATTTGATTTACACATCAC 68723
QY      4856 CCAGCTCCATTTTATTAATTTCTTCTCTGATGTTTGAAGTTCGCGCATGATATAT 4915
      |||
      68722 CCAGCTCCATTTTATTAATTTCTTCTCTGATGTTTGAAGTTCGCGCATGATATAT 68663
QY      4916 ATTAGATCTCAGAAATCTTGAATTTCTTCTGTTCTTAATCTTTAAAAATCATGTTTCC 4975
      |||
      68662 ATTAGATCTCAGAAATCTTGAATTTCTTCTGTTCTTAATCTTTAAAAATCATGTTTCC 68603
QY      4976 TTAATCTATCTTTTCTTATATTTGT 5000
      |||
      68602 TTAATCTATCTTTTCTTATATTTGT 68578

```

```

RESULT 4
AC116995 173131 bp DNA linear ROD 05-NOV-2003
LOCUS Mus musculus BAC clone RP23-3C16 from 1, complete sequence.
AC116995
AC116995.3 GI:21844700
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1. (bases 1 to 173131)

```

```

AUTHORS Holmes, A., Haglund, K. and Spalding, L.
TITLE The sequence of Mus musculus BAC clone RP23-3C16
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 173131)
AUTHORS Wilson, R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 173131)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Sequencing Center, 4444 Forest Park
4 (bases 1 to 173131)
REFERENCE 4 (bases 1 to 173131)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2002) Genome Sequencing Center, 4444 Forest Park
5 (bases 1 to 173131)
REFERENCE 5 (bases 1 to 173131)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2002) Genome Sequencing Center, 4444 Forest Park
6 (bases 1 to 173131)
REFERENCE 6 (bases 1 to 173131)
AUTHORS Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 16, 2002 this sequence version replaced gi:20389753.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
-----
Summary Statistics
Center project name: M_BA003C16
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RP23-3 BAC Library has been constructed by Kazutoyo Osagawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES

source

```

1.173131
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"

```



```
/clone="RP23-3C16"
repeat_region /tpc_family="Alu" 1084..1097
repeat_region /tpc_family="B4" 1098..1194
repeat_region /tpc_family="ID" 1643..1708
repeat_region /tpc_family="B2" 2034..2233
repeat_region /tpc_family="B2" 2401..2762
repeat_region /tpc_family="MaLR" 3195..3410
repeat_region /tpc_family="L2" 3450..3544
repeat_region /tpc_family="MaLR" 4140..4860
repeat_region /tpc_family="L1" 4837..4891
repeat_region /tpc_family="ERV1" 5137..5192
repeat_region /tpc_family="L1" 5594..5658
repeat_region /tpc_family="ERV1" 6208..6606
repeat_region /tpc_family="ERVK" 6754..7097
repeat_region /tpc_family="ERVK" 7191..7410
repeat_region /tpc_family="L1" 7415..7513
repeat_region /tpc_family="L1" 9015..9225
repeat_region /tpc_family="B2" 9683..9836
repeat_region /tpc_family="B4" 9708..9843
repeat_region /tpc_family="Alu" 9876..9990
repeat_region /tpc_family="B2" 10126..10319
repeat_region /tpc_family="L2" 10456..10637
repeat_region /tpc_family="B2" 12237..12655
repeat_region /tpc_family="L1" 12648..13442
repeat_region /tpc_family="L1" 13445..13614
repeat_region /tpc_family="ERV" 13452..13585
repeat_region /tpc_family="ERVK" 14087..14294
repeat_region /tpc_family="B4" 14257..14624
repeat_region /tpc_family="MaLR" 14622..14689
repeat_region /tpc_family="B4" 17030..17193
repeat_region /tpc_family="B4" 17231..17277
repeat_region /tpc_family="B4" 18109..18231
repeat_region /tpc_family="B4" 18757..18982
repeat_region /tpc_family="MaLR" 19152..19291
repeat_region /tpc_family="Alu" 19667..19876
repeat_region /tpc_family="B4" 20109..20258
repeat_region /tpc_family="B4" 20984..21106

repeat_region /tpc_family="Alu" 24485..25438
repeat_region /tpc_family="L1" 25873..25992
repeat_region /tpc_family="L1" 26010..26117
repeat_region /tpc_family="B4" 26042..26154
repeat_region /tpc_family="Alu" 27770..28295
repeat_region /tpc_family="ERV1" 29705..30153
repeat_region /tpc_family="L1" 30464..30599
repeat_region /tpc_family="Alu" 31056..31202
repeat_region /tpc_family="MER1_type" 32828..32936
repeat_region /tpc_family="Alu" 32855..32961
repeat_region /tpc_family="B4" 32988..33382
repeat_region /tpc_family="MaLR" 33454..33643
repeat_region /tpc_family="L1" 33651..33765
repeat_region /tpc_family="Alu" 33882..33984
repeat_region /tpc_family="Alu" 34576..34617
repeat_region /tpc_family="B2" 34628..34770
repeat_region /tpc_family="B2" 35256..35305
repeat_region /tpc_family="B2" 35531..35690
repeat_region /tpc_family="L1" 35708..35905
repeat_region /tpc_family="B2" 36174..36385
repeat_region /tpc_family="B2" 38346..38491
repeat_region /tpc_family="Alu" 38961..39070

Query Match 10.8%; Score 541.8; DB 10; Length 173131;
Best Local Similarity 68.3%; Pred. No. 1.5e-92;
Matches 977; Conservative 0; Mismatches 372; Indels 82; Gaps 13;

QY 1787 ATGTCAATCCCTAATACTATATATACCAATAGACACTTAATTGTGCCAAGCAACA--- 1843
DB 115748 ATGCACATCTCGGAAAGCGAGTATATCCACAGAACTGACTGTGCTGAGTGAAGAG 115807
QY 1844 ----ACCTTAACGTGGCCAAATTTATTTATTCATCAATAACAGCTGCTGTTTCACTTG 1899
DB 115808 ACCTTAACGTGAGTGGCCAGTTGATTTCT-CTCGGACAGACCTCTGTTGTTGAGTGG 115866
QY 1900 TGCACATCTGAATGAGCAATCCCTGTCGTGATGAGTCTTTCTTGCACTGATTAAGGANA 1959
DB 115867 TGCACACATGAGCGGTAGC--TGCATGCGGTTCTGATGATTTCTTGCACTGATGAGCTTA 115924
QY 1960 AACTGCTGAAGTTTGAGGCTGCTCCAGGACAGAGCCATCATGTGATGATCAATGAAAGCTTC 2019
DB 115925 AACTGCTGAAGTCAATGAGACTGCTCCAGGACAGAGCCCTGCTGTGATGATCAAGAAAGCTTC 115984
QY 2020 CAGCGCTGACCTCTGGCAAAAAGAGAGAAAGATAGAGAGGAGGAGTGGGGGAAA 2079
DB 115985 CTCTCTGCTGACCTCTGCGAGAGGGGAGACAAA-----TGAAGAGAGCGGAGTGGAGAA 116039
QY 2080 GGTTCAAGTCGGGTTTCTCTTGAACCTACAGATTAATGGGTCAAGAGCTGTGTGCAA 2139
DB 116040 GGTTCAAGGCGGGTTTCTCTTGAACCTGAGTATGGGTCAAGAGCTGTGTGCAA 116099
```

QY 2140 AGACTGTACAGCTGGATGACGCTGCTACCATTTGTCAGAGGAGGCGACACAGCT 2199
 DB 116100 AGAGGCTACAGCCAGGATGCAAGTTGCTACCCGTTGAGAGGAGGCGACAGCT 116159
 QY 2200 GTCCGAGGAGTCAAGCTGAGAGACCAAGGAG--GCTCGTTCAAGGATTTGATTCACAG 2256
 DB 116160 GTCCAGAGAACCCAGCTGAGAGACCAAGGAGCTGCTATTCAGAGTACTGATCCCGG 116219
 QY 2257 AGAGAGCATTTTCTGTATTTGATTAACCAAGAGATTCAAGACTC--CCTTTTGTAGC 2313
 DB 116220 AG-GAGCATCTTTCTTATTTAGTAACCAAGCAATTCAGACAGACACTCCCTTTTGTAC 116278
 QY 2314 GGGATCTGATTTCTTCCGCTAGCTGTAACCAATTAATTAATTTCAATTAAGTCAC 2373
 DB 116279 AGAGCTTTATTTCTTCCGCTAGCTGTAACCAATTAATTAATTTCAATTAAGTCAC 116338
 QY 2374 AGAAATTTATGCTGTATTTATCAATTTGGGGAATTTCTGTAAACCAAGGAGGAA 2433
 DB 116339 AGAGATTTATGCTGTATTTATCAAGTTTGGGGAATTTCTGTAAACCAAGGAGGGA 116398
 QY 2434 ATAACTCTTGGCTTTGGGCTGACGAATCTCACTTGGCTTAAATGTCAGAAAGTAG--T 2490
 DB 116399 ATAACTCTTGGCTTTGGG--CCGCTTTTGAATTTAGAAATGAGAAAGTGGTTCT 116448
 QY 2491 TCTCTCAAAATCTCTAAAGTCTTAATTTAAGAGCTGAAACTTAAGGAGAGCTGCAGT 2550
 DB 116449 TCTCTGACCTCTGAGCCACTTAATTTGAGGAGCTGAACTGTAAGGAGAGCTGCAGT 116508
 QY 2551 ATTATGTTGATGCTATGATTTGAAACTTAATTTAGTTATGATTTATTTAGCATGC 2610
 DB 116509 ACCGTTAGTCCGCTGTGATTTCAAGTTTCAATTAATTTAGCTGTATTTAGCATGAT 116568
 QY 2611 CATATGATTTCCCTTACAGCAATTAATTTAGTGACATGAAAAAAGCCAGACTTA 2670
 DB 116569 CTATCTAAATTTATTTAAGGCAATTTAATTAAGTCAAGAGAGTGGGCGATACGAGG 116628
 QY 2671 AACAGAAAAAGTTGCAAAATCATCAAGAGATTAGTTAACTGAAATGTTAAAG 2730
 DB 116629 AAATCTGATTAATTAAGAAATTTGGAGACACCTT-----AAGAGACTGGGT 116678
 QY 2731 CACATTTTATGTTGAGAAAGAAATTTGATTTTCAAGAGTTGATACCATTTATGCTTTT 2790
 DB 116679 AAATCTGATGTTGAGAAAGAAATTTGATTTTCAAGTAATTTGATACCATGTTGCTCTA 116738
 QY 2791 TCAGGAT--CTTTCAAGAAAAAGCTTTTGGGCGATACAGAAAGCTTTAGAAAAATTTG 2848
 DB 116739 AGAGATCTGATTTTCAAGAAAAAGCTTTTCAAGGAGTGTGGAAAGCTTTACAAACATCAA 116798
 QY 2849 AAGAGTAAATGAGCAAAATTAAGAAA-----AAATGT 2883
 DB 116799 ACCAAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAGT 116858
 QY 2884 TTTACAGGACATGATCTTTTACCTTTGATTAATTTATTTTGTCTTTCTTTTCTC 2943
 DB 116859 TTTGCGCCATGATCATTTGATCTTTTACCATTTTGTCTCTGTAAGTTCTTT----- 116912
 QY 2944 TAGTTAACAAGCTTAAAGAAAGCATCTATTTTCCATGATTTTACCCCAAGGAGACTT 3003
 DB 116913 ----TCTCTAGCTTAAAGAAAGCATCTGATTTCCAGGCTTATTTCCCAAGGAGACTG 116968
 QY 3004 CAAGCTGCGCTGCGCAAGTCAAGAGTTCAAGAGTTAAAGGCAAGGAGTACAGATTTGTTT 3063
 DB 116969 GAGGCTGCGCTGCGCAAGTCAAGAGTTAAAGAGTTAAAGGCGCGGAGTAAACATGATTTCTTT 117028
 QY 3064 CAGAGAGAGATCTGAGGAGGCTTTTCAAGTGTGATGAGAGCTGATGTTGATCAAG 3123
 DB 117029 CAGAGAGAGATCTGAGGAGGCTTTTCAAGTGTGATGAGAGCTGATGATGATCAAG 117088
 QY 3124 AAGAAATTTATGATGCTGTCTACAGAGCGGCGCTTTGTTCCAGAGCA 3174
 DB 117089 GAGAAATTTATGATGCTGTCTACAGAGCGGCGCTTTCCAGAGCA 117139

RESULT 5
 AC128908
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-320D4, WORKING DRAFT SEQUENCE, 2
 unorderded pieces.
 AC128908
 AC128908.3 GI:25073564
 HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_FULFILLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 179341)

Muzny, D., Marie, M., Merker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denon, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Grebergojz, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
 Hollins, B., Howell, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lounsbury, L., Louisse, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Matlines, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munadas, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwakoelameli, O., Okunolu, G., Olarunmugbon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
 Piopier, F., Poindexter, A., Popovic, D., Primus, E., Fu, L., L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Register, M., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartbeyn, A., Sison, I., Sitter, C. D., Smaj, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Stemle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Uman, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, J., Weller, X., White, F.,
 Williams, G., Wilson, R., Wleczek, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, D., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 179341)
 Worley, K. C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 179341)
 Rat Genome Sequencing Consortium.

REFERENCE
 AUTHORS

TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 19, 2002 this sequence version replaced gi:23911240. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBDA
Center clone name: CH230-320D4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 151376 bases at least Q40
Consensus quality: 153458 bases at least Q30
Consensus quality: 155000 bases at least Q20
Estimated insert size: 152913; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 178124: contig of 178124 bp in length
* 178125 178224: gap of unknown length
* 178225 179341: contig of 1117 bp in length.
Location/Qualifiers
1. 179341
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-320D4"
1. 2063
/note="wgs contig"
misc_feature 2565..3831
/note="wgs contig"
misc_feature 5301..6700
/note="wgs contig"
complement(176332..177098)
/note="clone boundary
clone_end:Sp6
site:
end_sequence:B2171386"

ORIGIN
Query Match 10.0%; Score 498.6; DB 2; Length 179341;
Best Local Similarity 66.5%; Pred. No. 2.2e-84;
Matches 921; Conservative 0; Mismatches 379; Indels 85; Gaps 11;
QY 1840 AAGAACTTAAGTGGCAAAATTTTATTTATTAATCAATACAGCTCGCTTTTCAGTGG 1899
DB 20604 AAGAGCTCTGTGTGACCAAGTTATTTCTCGGTAGAGCTCTTGTTCGGGTG 20663

QY 1900 TGCATCTGAATGCAAG-CAATCCCTGTCTGATGTGAGCTTTCTGCATGATAAGAA 1958
DB 20664 TGCACACATGAGCGGCAGCCGCTGGTTCGATCATAGTCTTTGCACTGATAGCTTA 20723
QY 1959 AAACGTCTGAAGTTGTGAGGCTGCTCCAGCGAGACCCATCATGTGATCATATGAAGCT 2018
DB 20724 AAACGTCTGAAGTTGTGAGGCTGCTCCAGCGAGACCCCTCGTGTATGATCAGAAAGCT 20783
QY 2019 CCAGCTGTGACCTCTGCA-AAAAAGGAGAAACAAGATAGAGAGCGAGTGGGCA 2077
DB 20784 CCACTCTGTGACCTCTGCAAGAGGGGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAG 20843
QY 2078 AAGTTTCAAGTGGGTTTCTCTTGAACCTCAAGATATATGTTTCAAGAGCTGTGTC 2137
DB 20844 GAGGTTCAAGTGGGTTTCTCTTGAACCTCAAGATATATGTTTCAAGAGCTGTGTC 20903
QY 2138 AAGACTGTACAGCTGTGAGTGCAGCTGTACCATGTTTCAGAGGAGGAGGACACAG 2197
DB 20904 AAGAGCCGCGAGCCAGATGCAAGTGTCTACCCCTTCGAGGAGGAGGAGGAGGAGGAG 20963
QY 2198 CTGTGAGAGAGTCAAGCTGTGAGACCAAGAG- -GCTGCGTTCAAGTATTTGTATC 2251
DB 20964 CTGTGAGAGAGTCAAGCTGTGAGACCAAGAGCTGTGAGTATTTGAGTCTGTATC 21023
QY 2252 CCAGAGAGAGATCTTTCTCTATTTATTAACCAAGAGTTTACATCTCCT- -TTT 2308
DB 21024 CCGGAG- -GAGATCTTTAG- -TAAACCAAGCAATTCAGACACATCTCCTCTT 21075
QY 2309 GTAGCGGAGTCTGATTTCTTCTGCGTAGGTCTTAAACCAATTAAGAAATTTCTATTAA 2368
DB 21076 GTAGAGGAGTCTGATTTCTTCTGCGTAGGTCTTAAACCAATTAAGAAATTTCTATTAA 21135
QY 2369 GTACAGAAATTTATGCTGTAGTTATCAATTTGGGAAATTTCTTTAAACCAAAAG 2428
DB 21136 GTACAGAGAGATTTATGATGTATTAACAGTTTGGGAGTAGTCTTTAAACAGAGCA 21195
QY 2429 GAAAAATATCTTTGGCTTTGGGCTGACAGAACTCCTGCTTGAAGTGAAGAAAGTA 2488
DB 21196 GGGGAATAGACCTTTGGGTTTG- - - - -CTCTTTTGAATTTGAAACAGAGAAAGTG 21245
QY 2489 G- - -TTCTCTCAAAATCTCTAAGTCTTAATTTACAGAGCTGAAACTTAAAGGAGAGCT 2545
DB 21246 GTTCTTCTGACATCTCTGAGCACTTAATTTAGAGGCTGACCTGTAGAGGAGAGCC 21305
QY 2546 GCAGTATTAGTTGTATGCTATGATTTGAACTTTAGTAAATTAGTATGATTTAGC 2605
DB 21306 GCATATCGGTGTGCTGCACTGTGATTCAGTTTCAGTAATTTAGCTGTGATTCATAGC 21365
QY 2606 AATGCTATGATTTATCCCTTACAGCAATTAATTAAGTGACATGAAAAAAGCCAG 2665
DB 21366 AATGCTATGATTTATCCCTTACAGCAATTAATTAAGTGACATGAAAAAAGCCAG 21425
QY 2666 ACTTAACAGAAAAAAGTTGCAAAATCATCAAGAGATTTAGTTAACTGAATGT 2725
DB 21426 GGGGCAAACTGATTAATTAAGGAAATGAAATTAACCTCAAG- - - - -AGACT 21475
QY 2726 AAGACATTTTATGAGTGAAGAAAGATTTTCAAGAGTTGATACCATTTAGGT 2785
DB 21476 GGGTAACTGTGAGAGTGAAGAAAGATTTTCAAGAGTTGATACCATTTAGGT 21535
QY 2786 CTTTTCAGGAGT- - -CTTTCAAGAAAGTCCCTTTGGGGGTACAGAGACCTTAGAA 2841
DB 21536 TTGTCTTAAGATCTGACCTTTCAGAAACCTGCTTTATGTTAGAGAGCTTACCAACA 21595
QY 2842 ACATTGAAGAGTGAAGTGAAGCAAAATTAAGAAATTAAGTTTACAGGAGCTGAAT- 2900
DB 21596 ACAACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 21655
QY 2901 - - - - -CTTACTTTCATTAATTTAT 2922
DB 21656 CCAGCAATTAAGAAAGCAAGAAAGGTTTGCAGGCAAGAGTGAATGACATTTGT 21715
QY 2923 TTCTGCTCTTTCTTTTCTAGCTAACAGACCTTAAGAAAGAGCATCATTTCTCATGACT 2982

```

Db      21716 GTCCGCTTCTCTCTCCAGCTTAACAGCTTAAAGAACATCATGTTCTCGGCT 21775
Qy      2983 TCATTACCCCCAGGGAGCTTCAAGCTGCCCGCCAGCAAGTTCAAGATAAAGGCCA 3042
Db      21776 TCATTACCCCCAGGGAGCTTCCAGGCTGCCAGAGTCCAGAAAGTTGAAGATAAAGGCCA 21835
Qy      3043 CGAGATPAACAGTATGTTTTCAGAAAGAGATCTGAGCGAGGCTTTTCAGTGGTGAAT 3102
Db      21836 CGAGATPAACAGTATGTTTTCAGAAAGAGATCTGAGCGAGGCTTCCATGGTGAAT 21895
Qy      3103 GAGCAGCTGATGTTGATPAACAAGATTTAATGAGCTTTGTCTAGAGGCGCGCCCTT 3162
Db      21896 GAACCCCTGACAGGACATCAAGAGAAATTCATGAGTGCCTGTCTAGAGAAAGTACTAG 21955
Qy      3163 GCTTC 3167
Db      21956 CCTTC 21960

```

```

RESULT 6
LOCUS    GS0119          431 bp      DNA      linear      STS 30-MAR-2000
DEFINITION SHGC-83058 Human Homo sapiens Sns genomic, sequence tagged site.
ACCESSION GS0119
VERSION   GS0119.1 GI:5221446
KEYWORDS STS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 431)
AUTHORS   Olivier, M. and Cox, D.R.
TITLE     Unpublished, Olivier, M., Cox, D.R. (2000)
JOURNAL   Unpublished (2000)
COMMENT

```

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: CCTCCCTTAATCAATGACG
Primer B: TCCAACTTCCACCTTGACTTA
STS size: 290
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul
Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

BAC ends sequenced at TIGR from the RPC11 BAC library. Designed and developed at the Stanford Human Genome Center.

FEATURES
Source
1. .431
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/map="2"
/clone_lib="Human"
STS
primer_bind 137..426
primer_bind 137..159
complement (404..426)
ORIGIN

Query Match 8.0%; Score 401.4; DB 11; Length 431;
Best Local Similarity 98.5%; Pred. No. 8.4e-66;
Matches 405; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Qy      3647 ATTCAGAAATTCAAAGAAATAGACATTCCTCTACTGACCCCAAGAACATATAC 3706
Db      21 ATCCGGAAATTCAAAGAAATAGACATTCCTCTACTGACCCCAAGAACATATAC 80
Qy      3707 TCTTCAGGCTGTGGAGGACACAGTTGGTAAGCGTCTTAACAGTTTTTTATATCCCT 3766
Db      81 TCTTCAGGCTGTGGAGGACACAGTTGGTAAGCGTCTTAACAGTTTTTTATATCCCT 140
Qy      3767 CCTTAATCAATACAGAGTTTGTATAGGCAACCTGAAATTCGCTTCATTCCTC 3826
Db      141 CCTTAATCAATACAGAGTTTGTATAGGCAACCTGAAATTCGCTTCATTCCTC 200
Qy      3827 CACCTGCGCTTATAGAAAGAACTGAAGTTGGTTCTGCAATATATGTATGCAAAAG 3886
Db      201 CACCTGCGCTTATAGAAAGAACTGAAGTTGGTTCTGCAATATATGTATGCAAAAG 260
Qy      3887 ATGATPAATCTAGATTTTATATATTGCAAAATACAAATATCTGGAATTAAT 3946
Db      261 ATGATPAATCTAGATTTTATATATTGCAAAATACAAATATCTGGAATTAAT 320
Qy      3947 ACTGCTTATCCAAAGCTAAGTAAATTTGGTAAACAACTTGTAAATATATG 4006
Db      321 ACTGCTTATCCAAAGCTAAGTAAATTTGGTAAACAACTTGTAAATATATG 380
Qy      4007 TAAAGATCCATGAATCCCTTTTATGCAAGTGGGAAAGTTGATGTC 4057
Db      381 TAAAGATCCATGAATCCCTTTTATGCAAGTGGGAAAGTTGATGTC 431

```

```

RESULT 7
LOCUS    AL451085          182166 bp      DNA      linear      PRI 09-MAR-2002
DEFINITION Human DNA sequence from clone Rp11-307C12 on chromosome 1, complete
ACCESSION AL451085 AC027440
VERSION   AL451085.20 GI:19309454
KEYWORDS HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 182166)
AUTHORS   Brown, A.
TITLE     Direct Submission
JOURNAL   Submitted (09-MAR-2002) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           humquerry@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
           On Mar 8, 2002 this sequence version replaced gi:19031396.
           During sequence assembly data is compared from overlapping clones.
           When differences are found these are annotated as variations
           together with a note of the overlapping clone name. Note that the
           variation annotation may not be found in the sequence submission
           corresponding to the overlapping clone, as we submit sequences with
           only a small overlap as described above.
           This sequence was finished as follows unless otherwise noted: all
           regions were either double-stranded or sequenced with an alternate
           chemistry or covered by high quality data (i.e., phred quality >=
           30); an attempt was made to resolve all sequencing problems, such
           as compressions and repeats; all regions were covered by at least
           one plasmid subclone or more than one M13 subclone; and the
           assembly was confirmed by restriction digest. The following
           abbreviations are used to associate primary accession numbers given

```


Db	5348	AGAGCGACGACC	GGGTCTCAAAAAAAAAAAAAAAAAA	ATCTAGAAATGGACGACTAGTATGA	5289				
Qy	450	ACTT--GACCTGGGACAT	TAACCTTTATGTGTATGAATTCACATCTTTAGAGAA	GAAT	507				
Db	5288	TTGTACAAACAT	TATGAAAGTACTTAGTGGCCCAACCTGATATCTTAAAA	TGGTTCAT	5229				
Qy	508	AGCATTTCTGAT	TAATAATGATATATATATATATTTATTTATTAATTC	AAATTCGAATTAAT	567				
Db	5228	TTTATGTTATGTA	TATGTATACCAATTAATTAATTAATTAACAGCAGCT	TACATTTATGTT	5169				
Qy	568	TCGTAGAAAC	TAGCTTCTCACTCTCTGATGTCAGTCAAAA	CTTTATGTTCTTTGGCC	627				
Db	5168	GC-----	TTTTATGATCATGAGATATATCTA	AAAACTTTAGCTATGTGGCC	5122				
Qy	628	GGGTGGCGGTGCTCA	CGCTGTAAATCCACAGACTTTTGAGAGCGCGAGCGGGTGA	ATCAC	687				
Db	5121	AGGAGCGGTGGCTCA	CGCTGTAAATCCACAGACTTTTGAGAGCGCGAGCGGGTGA	ATCAC	5062				
Qy	688	AAGGTTAGAGATG	AGAACCAATCCTGTGCTAACAGGAGAAACCTCGTCTCA	TAAAT	747				
Db	5061	GAGGTACAGGAATG	AGAACCAATCCTGTGCTAACAGGAGAAACCTCGTCTCA	TAAAT	5002				
Qy	748	ACAAAAAATTTAG	CGGGTGGCGGTGTCACAGCGCTGTATGCCACTGCTCAGAGGCTGA	807					
Db	5001	ACAAAAAATTTAG	CGGGTGGCGGTGTCACAGCGCTGTATGCCACTGCTCAGAGGCTGA	4942					
Qy	808	GGCAGAGAAATG	TGTGAACCCGGGAGCGGAGCTTGCATGAGCCGAGATTTGCGCCACT	867					
Db	4941	GGGAGAGAAATG	GCGTAAACCCGGGAGCGGAGCTTGCATGAGCCGAGATTTGCGCCACT	4882					
Qy	868	GCACTCCAGCT	TGGGCGACAGTGGGAGACTCTGTCTCA	AAAAAAAAAAAAAAAAAAAA	924				
Db	4881	GCACTCCAGCT	TGGGTTACAGAGGAACTCTGTCTCA	AAAAAAAAAAAAAAAAAAAA	4825				
RESULT 10									
LOCUS	AC008753	147330 bp DNA linear PRI 22-MAR-2003							
DEFINITION	Homo sapiens chromosome 19 clone CTD-310226, complete sequence.								
ACCESSION	AC008753								
VERSION	AC008753.9	GI:25140998							
KEYWORDS	HTG.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Mullalyrota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumariata; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE	1 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.								
TITLE	Direct Submission								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute.								
TITLE	Direct Submission								
JOURNAL	Submitted (03-APR-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA								
REFERENCE	3 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.								
TITLE	Direct Submission								
JOURNAL	Submitted (23-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA								
REFERENCE	4 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.								
TITLE	Direct Submission								
JOURNAL	Submitted (06-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA								
REFERENCE	5 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.								
TITLE	Direct Submission								
JOURNAL	Submitted (21-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA								
REFERENCE	6 (bases 1 to 147330)								
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.								
TITLE	Direct Submission								

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

COMMENT

Thomas, S., Uemami, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 175826)

Worley, K.C.

Submitted (22-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2000 this sequence version replaced gi:9719697.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: HMZB

Center clone name: RP11-498A2

Sequencing vector: M13: L08821

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 144734 bases at least Q40

Consensus quality: 162873 bases at least Q30

Consensus quality: 169424 bases at least Q20

Estimated insert size: 170770; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 3.3x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 22 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 21031: contig of 21031 bp in length
21032 21131: gap of unknown length
21132 40475: contig of 19344 bp in length
40476 40575: gap of unknown length
40576 57617: contig of 17042 bp in length
57618 57717: gap of unknown length
57718 75708: contig of 17991 bp in length
75709 75808: gap of unknown length
75809 89018: contig of 13210 bp in length
89019 89118: gap of unknown length
89119 100348: contig of 11230 bp in length
100349 100448: gap of unknown length
100449 110704: contig of 10256 bp in length
110705 110804: gap of unknown length
110805 123156: contig of 12352 bp in length
123157 123256: gap of unknown length
123257 132467: contig of 9211 bp in length
132468 132567: gap of unknown length
132568 141461: contig of 8894 bp in length
141462 141561: gap of unknown length
141562 146576: contig of 5015 bp in length
146577 146676: gap of unknown length
146677 151201: contig of 4525 bp in length
151202 151301: gap of unknown length
151302 155722: contig of 4421 bp in length
155723 155822: gap of unknown length
155823 160148: contig of 4326 bp in length
160149 160248: gap of unknown length
160249 163024: contig of 2776 bp in length
163025 163124: gap of unknown length

FEATURES

source

1. 175826

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="3"

/clone="RP11-498A2"

ORIGIN

Query Match 7.6%; Score 378.2; DB 2; Length 175826;

Best Local Similarity 67.0%; Pred. No. 1.4e-61;

Matches 665; Conservative 0; Mismatches 303; Indels 25; Gaps 8;

5 GGCCATGCTGTGCTCAAGCGCTGATTCGCCAGACTTTGGAGGC-----AGAGAT 58
161959 GGCACACGCGGCGCTCAACCTGATCCAGCAATTTGGAGGCTAAGAGCGCAGAT 162018
59 CGCTGAGCTCAGGAATTGGAGACAAGCTAAGTAACTAGGAACCTGCTGTACA 118
162019 CACTTGAAGTCAGGAGTTGAGACACAGCTGCCCAACATGGAAATCTCCCTCTATT 162078
119 AATATTAAGAAATTTT--CCAGCATGCTGCGTGCACCCCGATGCGAGCTATTGGG 176
162079 AAAAAACAATAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 162138
177 AGCTGAGGTAGAGGAATGCTTGAAGCCAGAGATTGAAGACGCTTAGGCAACTAGT 236
162139 AGCTGAGGTAGGAGTGAATCACTGAGTCAGAACTTGAACCAAGCTGCGCAAGATGCT 162198
237 GAGACCTGTGTCTA-TAAAAAATAATAGCTGTGCTTGGCAGAGCCCTGCA--GGT 293
162199 GAAACCCATCTCTACTAATAAACAAATTAAGCGGCGTGGTGCACAGCCGTATATCC 162258
294 AGCTACTCGAAGACTGAGGTGGAGATCAC-TGAGCCAGAGAGGCTGAGCTGAGTG 352
162259 AGCTACTCGAAGAGCTGAGGAGGAGGAATCACTTAACCTGGAAGGAGAGGTTGAGTG 162318
353 AACAGTATCACCCAGCTGATTCAGGCTGGAAGAC-AGAGGAGAGCCCTGTTCCAAA 411
162319 AGCTGATATCACGCTAATGCACTTCAGCTGGGCAACAAGATGAATCTGTCTCAAAA 162378
412 AAAAAAAAAAAAAAAAAAATGCAAGAAAAGACATCTAACTTGACCTGGACATTACTT 471
162379 AAAAAAAAAAATAATTAAGCGGCGGTGTGTGCTGTATTCCTGTAATCAGAGAGCT 162438
472 TATGTATGAATTAACAATCTTTTAGAAGAAATTAGCATTTGATTAATGATATAT 531
162439 GAGGTGAGAGATCCCTTGAACCCGAGAGGAGAGAGTGAAGTGAAGTCAACACA 162498
532 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 581
162499 CTGCACTCAGCTGGGCGCAGAAAGCTGTCTCAATTAATAAAAAAAAAAATTCACAGA 162558
582 -TTTCACTCTCTCAGTTGATGATCAAAACCTTAATGCTTTGGCCGGGTGCGGTGCT 640
162559 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 162618
641 CAGGCTGTAATCCAGCACTTGGAGGCGGAGGCGGAGTATCAAGATTAGAGAT 700
162619 CAGGCTGTAATCAAGCACTTGGAGGCGGAGGCGGAGTATCAAGATTAGAGAT 162678

QY 701 CGAGACCATCTCGCTTAACAGGTGAACCTGCTCTACTAATAATAC-AAAAATTG 759
Db 162679 CGAGACCATCTCGCTTAACAGGTGAACCTGCTCTACTAATAATAAAAAATTG 162738
QY 760 CCGGCTGCGGTCCAGACCTGTAGTCCAGCTGCTCAGAGGCTGAGGACAGAAAG 819
Db 162739 CTGGGCAAGCGGTGGCCCTGTAGTCCAGCTACTCGGAGGCTTAAGGCAAGAAAG 162798
QY 820 GTGTGAACCCGGAGGCGAGAGCTTGCAGTGAGCCGAGATTGGCCACTGACCT 879
Db 162799 GGGTGAACCTCGGAGGCGAGAGCTTGCAGTGAGCCGAGATTGGCCACTGACCT 162858
QY 880 GGGCGACAGTGGAGAGCTGTCTCAAAAAAATAAAAAAATGAAATGCTTTG 939
Db 162859 AGCGCAGACAGGCGAGCTCGGCTCAAAAAAATAAAAAAATGAAATGCTTTG 162918
QY 940 AGCAAGTAGTCTCTCTTTCTCTCTCTTT 972
Db 162919 AGGATATTATCTCCACATTTTCTCTCTTT 162951

RESULT 12
AC136143 166148 bp DNA linear PRI 05-MAR-2003
LOCUS Papio anubis clone rp41-5m22, complete sequence.
DEFINITION AC136143
AC136143 GI:28850065
VERSION HTG.
KEYWORDS Papio anubis (olive baboon)
SOURCE Papio anubis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.

REFERENCE
AUTHORS Lau, C., Shaikh, T. and Roe, B. A.
TITLE 1* (bases 1 to 186148)
JOURNAL Papio anubis BAC Clone rp41-5m22
Unpublished
REFERENCE
AUTHORS Lau, C., Shaikh, T. and Roe, B. A.
TITLE 2* (bases 1 to 186148)
JOURNAL Papio anubis BAC Clone rp41-5m22
Unpublished
REFERENCE
AUTHORS Lau, C., Shaikh, T. and Roe, B. A.
TITLE Direct Substitution
JOURNAL Submitted (29-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Lau, C., Shaikh, T. and Roe, B. A.
TITLE Direct Substitution
JOURNAL Submitted (26-FEB-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Lau, C., Shaikh, T. and Roe, B. A.
TITLE Direct Substitution
JOURNAL Submitted (05-MAR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Mar 5, 2003 this sequence version replaced gi:28460812.

COMMENT
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR

FEATURES
source Location/Qualifiers

1. 186148
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="rp41-5m22"
/clone_1ib="RPCT - 41 Male (Olive) Baboon BAC Library"

ORIGIN
Query Match 7.5%; Score 377; DB 9; Length 186148;
Best Local Similarity 67.4%; Pred. No. 2.4e-61;

Matches 625; Conservative 0; Mismatches 285; Indels 17; Gaps 6;
QY 11 TCGTGGCTCACGGGTGATTTCCAGACATTTGGGAGGAG- - -GAGATCGCTGAG 66
Db 88926 TCGGTGGCTCACGGGTGATTTCCAGACATTTGGGAGGAGGAGGAGGAGATCAAG 88985
QY 67 CTCAGAAATTGGAGCAAGCCCTACGTAACATATGTAAGAACTCTGTCTGTAATAATA 126
Db 88986 GTGAGGAATTCAGACATCTCTGGCTTACCGGTGAAATCCGTTTCTACTAAAAATA 89045
QY 127 AAGAAATTTCCAGGAGTGGGCGTGACCCCGAGTGCAGCTATTTGGAGGCTGAGGT 186
Db 89046 AAAAATCTGGCTGGGCGCAATGGCTCAACCTGTAAATCCAGACATTTAGAGGCGAGGC 89105
QY 187 AGAGGAATGCTTGAAGCCAGAGATTGAAGAACAAAGCCTTAGGAACATAGTGAACCTGT 246
Db 89106 GGAAGAACAACTGAGGTCTAGAGATTGAGACACACCTGACCAACATGCAAGAAACCCAT 89165
QY 247 GTCTATAAAAAAT- - -AATTAGCTGTGTCTTGGACAGGCTGCA- -GCTAGCTACT 300
Db 89166 CTCTACTAAAAATACAAATTTAGTGGGATATGGCGATATGGCTGTAAATCCAGCTACT 89225
QY 301 CGAAGACTGAGGTGGAGGATCAC- TGAGCCAGAGGCTGAGGCTGCAATGAAACAGTG 359
Db 89226 CGGAGGCTGAGGCGGAGAAATCGCTTGAACTGGAGGCGAGGTTGCAAGAGCCGAG 89285
QY 360 ATCACCAGCTGAGATTCACGCTGGAAGACAGAGGAGACCTGTGTTCCAAAAAATA- - 417
Db 89286 ATCGACCTTTGACCTCCAGCTTGGCAACAGACCAAAATCCATCTCAAAAAAATAAG 89345
QY 418 AAAAAAATAATGCAAGAAAGACATCAATCAATGACCTGCAATTAATTTATG 477
Db 89346 ATAAATTAATAAACTAATCTAATTAATTAATTAATGATCGGCGGTGTGGCGCATGCG 89405
QY 478 ATGAATATTCACATCTTTTGAAGAGAAATTAGCAATTTGATTAATATGATATAT 537
Db 89406 TGTAAATCCAGCTACTTTAGAGGTGAGGAGAGAAATGATGTGAACCTGGAGAGCGAG 89465
QY 538 ATTATTAATTAATTCAAATGGAATTAATATTTGAGAAATAGCTTCTCACTCTCAT 597
Db 89466 ATTCAGATGAGCGGAGATGCACTGCACTCCAGCTGGGCGATAGATGAGACTCGG 89525
QY 598 TGTCACTCAAACTTTAATGATGCTTTGGCGGAGTGGGCTCAGCCTGTAAATCCAG 657
Db 89526 TCTCAAAAAA- - -AAAAAATAAGGCGGCGGAGGCTCAAGCCTGTAAATCCAG 89581
QY 658 CACTTGGAGGCGGAGGCGGAGTGAATCAAGATTAGAGATGAGACCATCTTGCTA 717
Db 89582 CACTTGGAGGCGGAGGCGGAGTGAATCAAGATTAGAGATGAGACCATCTTGCTA 89641
QY 718 ACAAGGTGAACCTGTCTCTAATAAATAAATAAATAATTAGCCGGGTGGGTCCAGAC 777
Db 89642 ACACAGTGAACCCGCTCTCTAATAAATAAATAAATAATTAGCCGGGTGGGTGGGTCC 89701
QY 778 GCTGTAGTCCAGGCTCAGAGGCTGAGGAGGAGATGATGTAAGCCGGAGGCG 837
Db 89702 GCTGTAGTCCAGGCTCAGAGGCTGAGGAGGAGATGATGTAAGCCGGAGGCG 89761
QY 838 GAGCTTCAGTGAAGCGAGATTGCGCACTGCACTCGGCGAGCAGTGGCGAGTGGAGACT 897
Db 89762 GAGCTTCAGTGAAGCGAGATTGCGCACTGCACTCGGCGAGCAGTGGCGAGTGGAGACT 89821
QY 898 CTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 924
Db 89822 CCGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 89848

RESULT 13
AP002906/c 162728 bp DNA linear PRI 03-APR-2001
LOCUS AP002906
DEFINITION Homo sapiens genomic DNA, chromosome 8q23, clone: KB189B1.
ACCESSION AP002906
VERSION AP002906.2 GI:13516382

Query Match	7.5%; Score 376; DB 9; Length 162728;
Best Local Similarity 60.8%; Pred. No. 3,7e-61;	
Matches 640; Conservative 0; Mismatches 265; Indels 25; Gaps 8;	
12 GCTGTGCTACGCGTGTATTCACAGACTTGGAGGCG-----GAGATCGCTTA 65	/rpt family="TTTTCCn" 18573 .18628 /evidence=not_experimental /rpt_family="TTCGCCn" complement(18631 .18667) /evidence=not_experimental /rpt_family="AluJo" complement(19046 .19344) /evidence=not_experimental /rpt_family="AluSp" 19436 .19456 /evidence=not_experimental /rpt_family="Ar Rich" complement(19457 .19516) /evidence=not_experimental /rpt_family="FLAM_C" 20700 .20991 /evidence=not_experimental /rpt_family="AluSx" 21256 .21373 /evidence=not_experimental /rpt_family="FLAM_C" 21729 .21846 /evidence=not_experimental /rpt_family="L2" 22372 .22496 /evidence=not_experimental /rpt_family="LtrR41" 22895 .22973 /evidence=not_experimental /rpt_family="LtrR50" 23184 .23207 /evidence=not_experimental /rpt_family="TGCn" 23412 .23514 /evidence=not_experimental /rpt_family="L2" 23616 .23708 /evidence=not_experimental /rpt_family="MIR" 27023 .27322 /evidence=not_experimental /rpt_family="AluDb" 27460 .27765 /evidence=not_experimental /rpt_family="AluSx" 28489 .28748 /evidence=not_experimental /rpt_family="AluSx" 28768 .28796
66 GCTCAGGAAATTTGAGACAGAGCTACGTATAGTGAACCTCTGTCTGTACAAATATA 125	
49273 GCTCAGGAGTTCAAGACCAAGCTGTGCAATATTTGTGAACC-TGTCTCTACAAAAATA 49215	
126 AAAAATTTTTCAGAGCATGTGTGGCTGACACCCAGTGGCACTATTTGGAGGCTGAAG 185	
49214 CATTAATTTGTCAGAGCGCGGTGCTCATGCTTGAATCCAGACTTTGGAGGCCAAG 49155	
186 TAGAGGAATCTTGAAGCAGAGTTGAAGACAGAGCTAGCAATATAGTGAACCTTG 245	
49154 CAGGCAATATCTGAGAGTGGAGTTCAAGACCAAGCTGACCAATATGAGAAATTTCCA 49099	
246 TGTCTATTAATAAT---AATTAGCTGTGTCTTGGCAAGGCTTGA--GCTAGCTAC 299	

Db	Accession	Source	Organism	Reference	Title	Journal	Authors
Db	49094	ACTCCACTAAAAATTACAAAATTAGCCGGGACATGCTGATGCTGTAATCCACGCTAC		49035			
Qy	300	TCGGAGACCTGAGGTGGAGGATCAC-TGAGCCACAGAGCTTGAGGCTGCAGTGAACACT		358			
Db	49034	TCGGAGAGCTGAGGCGAGAAATTACTTAACCTGGAGGCGGAGGTTGGCACTGAGCCCA		48975			
Qy	359	GATCACCCAGCTGGATTCCAGCCTGGAGACAGAGGAGACCCCTGTTCCAAAAAANA		418			
Db	48974	GATTGCCCATTCGACTCCAGCTGGGTGACA-AGAGCAAACTCTGTCTCAAAAAANA		48916			
Qy	419	AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGAGACATTACTTTATGTA		478			
Db	48915	AAAAAAAAAATCAATA-----ATTAGCAAGCGTGTGGCACATGCTTAAGTCA		48862			
Qy	479	TGAAATTCACAACTTTTAGGAAGAAATTAGCTTTCTGATAAATGTATTAATTATA		538			
Db	48861	GCTACTCGGAGGCTGAGGTGGAGGTTGGCTTAGAGCCAGAAATTGCACTGACTAGGA		48802			
Qy	539	TTATTATAAATTCGAAATGAAATTAATAATTTCTGAGAACTAGCTTCTCACTCTCAG--		596			
Db	48801	TCATGCCACTGCACTCCAGCTGGGGGACAGTCAACATGCTCAAAAGAAAAAGAG		48742			
Qy	597	--TTGTCAGTCAAAACTTTATATGTCTTTTGCCCGGTGGGTGGCTCAGCCCTGTAATCC		654			
Db	48741	AAGTTTCTTAAAAAGAGCAAAATCTCGGCCAGCGAGTGGCTACAGCCTGTATTC		48682			
Qy	655	CAGCACTTTGGAGAGCGGAGCGGAGTATCACAAGTTTAGAGATCGAACCATCTGG		714			
Db	48681	CAGCACTTTGGAGAGCTGAGCGGGTGTATCAGAGTCAAGAGATCGAGACCATCTGG		48622			
Qy	715	CTAACACGGTGAAACCTCGTCTCTACTAAAAATACAAAAAATTAGCCGGGTGGGCCA		774			
Db	48621	CTAACACGGTGAAACCCCATCTCTACTAAAAATACAAAAAATTATGTGTCAAGGGTGGG		48562			
Qy	775	GAGCCCTGTAGTCCAGCTGCTCAGAGGCTGAGGACAGAGAAATGTGTGAACCCGGAG		834			
Db	48561	GGCACCTGTATCCCACTACTCAGAGGCTGAGGACGAGAAATGTGTGAACCTGGAG		48502			
Qy	835	GCGAGGCTTGAGTGAAGCCGAGATTGGGCCACTGCACTCCAGCTGGGGGACAGTGGAG		894			
Db	48501	GTGAGGCTTGAGTGAAGCCGAGATCGAGCCACGCACTCCAGCCTGGGGACAGAGTGA		48442			
Qy	895	ACTCTGTCTCAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANA		924			
Db	48441	ACTCCACTCAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANA		48412			

Center code: CSHL
 Web site: <http://www.cshl.org/genseq>
 Contact: mccombie@cshl.org
 Project Information
 Center project name: 133K12
 Center clone name: 133K12

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 33563: contig of 33563 bp in length
 * 33564 34544: gap of unknown length
 * 34545 57280: contig of 22736 bp in length
 * 57281 58260: gap of unknown length
 * 58261 73887: contig of 15627 bp in length
 * 73888 74867: gap of unknown length
 * 74868 85395: contig of 11128 bp in length
 * 85396 86975: gap of unknown length
 * 86976 97534: contig of 10558 bp in length
 * 97534 98513: gap of unknown length
 * 98514 104746: contig of 6233 bp in length
 * 104747 105726: gap of unknown length
 * 105727 109390: contig of 3664 bp in length
 * 109391 110370: gap of unknown length
 * 110371 111469: contig of 1099 bp in length.

FEATURES

Source
 1. 111469
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p25"
 /clone="133K12"

ORIGIN

Query Match 7.5%; Score 374.2; DB 2; Length 111469;
 Best Local Similarity 68.1%; Pred. No. 8.3e-61;
 Matches 637; Conservative 0; Mismatches 263; Indels 35; Gaps 7;

5 GGCACATGCTGTGCTCAGCGCTGATTCACGACCTTTGG-----AGGAGAGAGAT 58
 51735 GGCAGGAGACAGTGGCCGATGCTGATTCAGACCTTTGGAGCGCTAAGGGGGTGGAT 51676
 59 CGCTTGAGCTCAGGAATTGAGACAGACCTTACGTAATGTAAGAACTCTGTCTGTACA 118
 51675 CACCTGAGGCTCTGGAGTTTGACACAGCTGCGCAAAACATGATGTAAC-CACCAATCTTACT 51617
 119 AATAATTAAGAAATTTCCAGGATGATGGGCGTGCAGCCCGCAGTGCACCTATTGGGAG 178
 51616 AAAAAATCAAAAAGTACCGGTCGATGCTGCGCCCTGTATCTTACACCTTTGGAG 51557
 179 GCTGAGTAGAGAGAAATGCTTAAGCCAGAGTGAAGCAAGCCCTAGGCAACATAGTGA 238
 51556 GCCGAGGAGCGGAATGCTTGAAGCATGTTGAACACAGCCCTGCGCAACATGCTGA 51497
 239 GACCTGTGTCTATAA-----AATATTTAGCTGTGTCTTGGCACAGGCTTGCA-- 290
 51496 AACCTTGTCTTACTAATAAACAACAAAATTTAGCTGGGGGTATGCGACATGCTGTAT 51437
 291 GCTAGCTACTCGAAGACTGAGGTGGAGAGATCAC-TGAGCCAGAGGCTGAGGCTGCA 349
 51436 CCCAGCTACTCGGCTGGCTGGGCAAGAAATCGCTTGAACCCGAGAGCGAGGTGGA 51377
 350 GTGAACAGTGAATCACCGACCTGATTCAGCCCTGGAAGACAGAGGAGACCTCTTTCCA 409
 51376 GTGAGCCGAGATCACGCCACTGCGCTTACCTGGGCGACAGAGCAAGACTCATCTCA 51317
 410 AAAAAAAAAAAAAAAAAATGCAGAAAAGACATCATTAACCTTGAACCTGGACATTA 469

DB 51316 AAAAAATTAATTAACAAAAT-----TAGCCAGCAAGTGGCAGCAC 51273
 QY 470 TTTATGATGATGAATTCACAAATCTTTAGAGAAATTAAGCATTTCTGATTAATGAT 529
 DB 51272 TGTACTCCAGCTACTTTGGAGCGGAGGACAGAAATCACTGTAACCTGGAGGAGA 51213
 QY 530 ATAATTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 589
 DB 51212 CTTCAGTAGAGCTGAGCGGACCACTCCTCAGACCGCCTGACAGAGCAAGACT 51153
 QY 590 CTTCAGTTGTAGCAAAACCTTTATAGTCTTTGGCGGGGGGGTGGCTGACGCTGT 649
 DB 51152 CTGTACAT--ACATPACAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 51096
 QY 650 AATCCAGCACTTTGGAGCGCGAGCGGCTGATCACAAGTTAGAGATGAGACCAT 709
 DB 51095 ATTCACAGCACTTTGGAGCGCGAGCGGCGAGATCACAAGTTAGAGATGAGATCAT 51036
 QY 710 CTTGGCTAACACGCTGAAACCTCGTCTTACTTAATAAATTAACAAAATTAATTAATTAAT 769
 DB 51035 CTTGGCTAACACGCTGAAACCTCATCTTACTTAATAAATTAACAAAATTAATTAATTAAT 50976
 QY 770 TGCCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGGAGAAATGATGTAACCC 829
 DB 50975 TGCCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGGAGAAATGATGTAACCC 50916
 QY 830 GGGAGCGGAGGCTTGCAGTGAAGCCAGATTTGGCCCATCTCAGCTCAGCTGGGCGCACT 889
 DB 50915 GGGAGCGGAGGCTTGCAGTGAAGCCAGATTTGGCTCAGTCACTCAGCTCAGCTGGGCGCACT 50856
 QY 890 GCGAGCTCTGTCTCAAAAAAAAAAAAAAAAAA 924
 DB 50855 GCGAGACTCCGCTCAAAAAAAAAATAATTAATTAATTAATTAATTAATTAATTAAT 50821

RESULT 15

AC018410 147024 bp DNA linear PRI 10-NOV-2002
 LOCUS Homo sapiens chromosome 11, clone RP11-17G12, complete sequence.
 DEFINITION AC018410
 ACCESSION AC018410.24 GI:24850147
 VERSION 1
 KEYWORDS HTG.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 147024)
 Birren, B., Nussbaum, C., Lander, E.
 Homo sapiens chromosome 11, clone RP11-17G12
 Unpublished

REFERENCE 2 (bases 1 to 147024)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckler, J., Bede, F.,
 Boguslavskiy, L., Boungkaler, B., Brown, A., Castle, A., Colangelo, M.,
 Collins, S., Collymore, A., Cooke, P., DeBellano, K., Dewar, K.,
 Domino, M., Doyle, M., Fenesfor, J., Ferreira, P., Fitzhugh, W.,
 Forrest, C., Gage, D., Galagan, J., Gardy, S., Grant, G., Hagos, B.,
 Heath, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L.,
 Karlas, A., Klein, J., Landers, T., Lebecky, J., Liu, C., Locke, K.,
 Macdonald, P., Margulis, N., McBwan, P., McElroy, A., McKernan, K.,
 Mellett, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T.,
 O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R.,
 Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J.,
 Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A., and Zody, M.

TITLE

Unpublished

AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckler, J., Bede, F., Boguslavskiy, L., Boungkaler, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeBellano, K., Dewar, K., Domino, M., Doyle, M., Fenesfor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardy, S., Grant, G., Hagos, B., Heath, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karlas, A., Klein, J., Landers, T., Lebecky, J., Liu, C., Locke, K., Macdonald, P., Margulis, N., McBwan, P., McElroy, A., McKernan, K., Mellett, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A., and Zody, M.

TITLE

Unpublished

AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckler, J., Bede, F., Boguslavskiy, L., Boungkaler, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeBellano, K., Dewar, K., Domino, M., Doyle, M., Fenesfor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardy, S., Grant, G., Hagos, B., Heath, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karlas, A., Klein, J., Landers, T., Lebecky, J., Liu, C., Locke, K., Macdonald, P., Margulis, N., McBwan, P., McElroy, A., McKernan, K., Mellett, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A., and Zody, M.

TITLE

Unpublished

AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckler, J., Bede, F., Boguslavskiy, L., Boungkaler, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeBellano, K., Dewar, K., Domino, M., Doyle, M., Fenesfor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardy, S., Grant, G., Hagos, B., Heath, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karlas, A., Klein, J., Landers, T., Lebecky, J., Liu, C., Locke, K., Macdonald, P., Margulis, N., McBwan, P., McElroy, A., McKernan, K., Mellett, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A., and Zody, M.

Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collimore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mhova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (04-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 147024)

REFERENCE
JOURNAL
AUTHORS
Birn, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collimore, A., Cooke, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mhova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (19-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 147024)

REFERENCE
JOURNAL
AUTHORS
Birn, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collimore, A., Cooke, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mhova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (10-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2002 this sequence version replaced g1:24137549.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

Center project name: L3636
Center clone name: 17_g_12

Only the last 147.0 kilobases of this clone are being submitted.
The remainder overlaps accession number AC090582 [WTCGR project L11783].

FEATURES
source
1. 147024
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-17G12"
/clone_1kb="RP11-11 Human Male BAC"
complement(1..60)
/rpt_family="Alu"
complement(100..193)
/rpt_family="L2"
345..372
/rpt_family="AT-rich"
426..732
/rpt_family="AluSg"
complement(733..999)
/rpt_family="AluO"
1033..1054
/rpt_family="AT-rich"
complement(1058..1286)
/rpt_family="MIR"
complement(12149..2459)
/rpt_family="AluY"
complement(2463..2753)
/rpt_family="AluSg"
complement(2786..2981)
/rpt_family="MIR"
2994..3151
/rpt_family="MIR"
complement(3975..4138)
/rpt_family="MER91A"
4557..4735
/rpt_family="TG)n"
4766..4950
/rpt_family="TG)n"
4996..5179
/rpt_family="TG)n"
5201..5313
/rpt_family="TG)n"
5339..5434
/rpt_family="TG)n"
5435..5727
/rpt_family="TG)n"
5728..5857
/rpt_family="AluSg"
5728..5857
/rpt_family="TG)n"
6048..6201
/rpt_family="TG)n"
6282..6458
/rpt_family="TG)n"
6477..6603
/rpt_family="TG)n"
6766..6942
/rpt_family="TG)n"
7001..7180
/rpt_family="TG)n"
7216..7288
/rpt_family="TG)n"
7305..7480
/rpt_family="TG)n"
7492..7611
/rpt_family="TG)n"
8199..8256
/rpt_family="C-rich"
8724..8751
repeat_region

```
repeat_region /rpt_family="ATG)n"
complement(8800..8932)
/rpt_family="MIR3"
repeat_region complement(9403..9710)
/rpt_family="AluUb"
repeat_region 9712..9788
/rpt_family="AT_rich"
repeat_region complement(9789..10086)
/rpt_family="AluUb"
repeat_region complement(10094..10269)
/rpt_family="MIR"
repeat_region complement(10275..10340)
/rpt_family="crRNA-Mec-1"
repeat_region 11030..11164
/rpt_family="MIR3"
repeat_region 11536..11720
/rpt_family="MIR"
repeat_region 12144..12212
/rpt_family="L2"
```

Query Match 7.5%; Score 374.2; DB 9; Length 147024;

Best Local Similarity 68.1%; Pred. No. 8.2e-61;

Matches 637; Conservative 0; Mismatches 263; Indels 35; Gaps 7;

```
QY 5 GGCCATGCTGTGGCTCAGCGCGTATTCGCCACTTGGG-----AGGAGAGAGAT 58
DB 79292 GGCCGGGACAGTGGCCCGCTGTATTCAGACACTTGGGCGTAAAGCGGTGAT 79351
QY 59 CGCTTAGCTCAGGAATTGGAGCAAGCCTAGTAACATAGTAAACCTGTCTGTACA 118
DB 79352 CACCTGAGGTCTGGAGTTGAGACAGCCTGGCAACATGTGAA-CAACATCTACT 79410
QY 119 AATAATAAAGATTTTCCAGGAGTGGCGCTGCACCCCAAGTGCAGCTATTGGAG 178
DB 79411 AAAAAACAAAAAGTAGCGGGTGCAGTGGCTGCAGCTGTATCTAGCACTTGGAG 79470
QY 179 GGTGAGTGAAGGATGCTTAAGCCAGAGGATGAAGCAAGCCTAGGCAACATAGTA 238
DB 79471 GCGAGGAGGAGGAAATGCTTAAGGATGAGTTGAAACAGCCTGGCAACATGTGA 79530
QY 239 GACCTGTGTCTATTA-----AATAATTAGCTGTGTCTTGGCACAGGCTGCA-- 290
DB 79531 AACCTGTCTACTATAAACAACAAAAATTAGCTGGGCTGATGCACTGTCTTAA 79590
QY 291 GCTAGCTACTCGAAGACTGAGGTGGAGATCAC-TGAGCCGAGAGGCTAGGCTGCA 349
DB 79591 CCCAGCTACTCGGCTGGCTGGGCAAGAAATCGCTTGAACCGAGAGGCGAGTTGCA 79650
QY 350 GTGAACAGTATCACCCAGCTGATTCGCTGGAAGACAGAGGAGACCTGTTCCA 409
DB 79651 GTGAGCCGAGATCAGCCACTGCGCTTGGCTGGGCGACAGAGCAAGCTCATCTCA 79710
QY 410 AAAAAAAAAAAAAAAAAAATGCAAAAAAGACATCATTAACCTTGACTGGGACATTA 469
DB 79711 AAAAAATATATATACAAAAT-----TAGCAGGCAAGTGGCAGCAC 79754
QY 470 TTTATGTGATGAATTCACAATCTTTTAGAAGAAATTAGCAATTCTGATTAATGAT 529
DB 79755 TGTAGTCCAGCTACTTGGGAGGCGGAGGAGAGAAATCACTTGAACCTGGAGGAGAA 79814
QY 530 AATAATTATTTATTAATTAATGAATGAATTAATATCTGAGAACTAGCTTCACT 589
DB 79815 CTGGAGTGAAGTGAAGCGGCACTGATCCAGACCGCGTGAAGAGAGCAAGACT 79874
QY 590 CTCTCAGTGTGAGTCAAACTTTAATGCTTTTGGCGGGTGGGCTGCTCAGGCTGT 649
DB 79875 CTGTCACT--ACATACAAAAATTAATAAGCCGGGCGGCTGCTCAGGCTGT 79931
QY 650 AATCCAGCACTTTGGGAGCGGAGCGGGTGGATCACAAGGTTAGAGATGAGACAT 709
DB 79932 ATTCTCAGCACTTTGGGAGCGGAGCGGGGAGATCACAAGGTTCAAGAGATGAGATCAT 79991
QY 710 CCTGGCTAACAGGGAACCTGTCTCTAATAAAATACAAAAAATTAGCCGGGTGGG 769
```

```
DB 79992 CCTGGCTAACAGGGAACCCCATCTCTAATAAAATACAAAAATTAGCCGGGCTGG 80051
QY 770 TCCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGACAGAGAAATGTGTGAACCC 829
DB 80052 TGGCGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGACAGAGAAATGGCATGAAGCC 80111
QY 830 GGGAGCGGAGCTTGCAGTGAAGCCGAGATTGGCCACTCCAGCCTGGGCGACAGT 889
DB 80112 GGGAGCGGAGCTGAGTGAAGCCGAGATTGCTCACTCCAGCCTGGGCTGACAGA 80171
QY 890 GCGAGCTGTCTCAAAAAAAAAAAAAAAAAA 924
DB 80172 GCGAGCTCCGTCTCAAAAAAAAAATTAATAA 80206
```

Search completed: December 27, 2004, 04:17:57

Job time : 13663.4 sec

This Page Blank (uspto)